



FEATURES	Location/Qualifiers									
	1..8589									
	/organism="unknown" /mol_type="unassigned DNA"									
ORIGIN										
Query Match	97.5%; Score 4184.8; DB 6; Length 8589;									
Best Local Similarity	98.7%; Pred. No. 0;									
Matches 4232; Conservative	0; Mismatches 47; Indels 9; Gaps 1;									
QY	3	GGTGGCGGACGAGCGCGCACTAGGCAAGCGTACGAGGAGGACCGCGCGACCGGACG	62							
DB	9	GGTGGCGGACGAGCGCGCACTAGGCAAGCGTACGAGGAGGACCGCGCGACCGGACG	68							
QY	63	GCTCGGGGACGAGCGCGCGCACTGGGACTGGGACGTCGACCGAGGCGTGGAGGCG	122							
DB	69	GCTCGGGGACGAGCGCGCGCGCACTGGGACTGGGACGTCGACCGAGGCGTGGAGGCG	128							
QY	123	GGGCGTGGGGCGCGGGTGGCGCTCCCGCGGTGCTGTCTCCACCGCTCGCGCCACGGCT	182							
DB	129	GGGACCGCGCGCGCGGTGCTGTCTCCCGGGGTGCTGTCCCGGGCGCTGCCCGCGCGC	188							
QY	183	GCTGCTGCTGCTGTGTGCTCCCGCGCGGTGCTGTCTCCACCGCTCGCGCCACGGCT	242							
DB	189	GCTGCTGCTGCTGCTGTGTGCTCCCGCGCGGTGCTGTCTCCACCGCTCGCGCCACGGCT	239							
QY	243	CGAGGCGCGGGCGCGCGCGGTGCTGCGGCTCAGCGCGAGCGCGGCGCAAGCAATG	302							
DB	240	CGAGGCGCGGTGCGGCGCGCGGTGCTGCGGCTCAGCGCGAGCGCGGCGCAAGCAATG	299							
QY	303	TGACGGCGCGTGTGCAAGCGGGTGTGCTGCAACCGTGGCACCGCGCGAGTGGCTGCGCC	362							
DB	300	TGACGGCGCGTGTGCTGCAAGCGGGTGTGCTGCAACCGTGGCACCGCGCGAGTGGCTGCGCC	359							
QY	363	CGCGCGTGGGTGGGCGCAATGCGACACTGCGGGGGCGCGCTTACAGACTAACTGGATC	422							
DB	360	CGCGCGTGGGTGGGCGCAATGCGACACTGCGGGGGCGCGCTTACAGACTAACTGGATC	419							
QY	423	TTCTGGGTTGTGACAGATGGACCTGGAAATATATAAATAAATAAATAAATAAATAAATAA	482							
DB	420	TTCTGGGTTGTGACAGATGGACCTGGAAATATATAAATAAATAAATAAATAAATAAATAA	479							
QY	483	CATTGAAGGACGACCAATATAGAAATATAGAGACTTCGTTTCAATCAATTTTGTACAGATG	542							
DB	480	CATTGAAGGACGACCAATATAGAAATATAGAGACTTCGTTTCAATCAATTTTGTACAGATG	539							
QY	543	TAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC	602							
DB	540	TAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC	599							
QY	603	ATTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGTTGTGC	662							
DB	600	ATTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGTTGTGC	659							
QY	663	CACATCAGGTTATGCTTGTGCAATTTTATGATGCTGCTTATATATTTGACTGGATT	722							
DB	660	CACATCAGGTTATGCTTGTGCAATTTTATGATGCTGCTTATATATTTGACTGGATT	719							
QY	723	TAATATTACTTACAGTTTGTATGTGTCGAAATAATGCTCAGGCGGAGGAGAGTGTAA	782							
DB	720	TAATATTACTTACAGTTTGTATGTGTCGAAATAATGCTCAGGCGGAGGAGAGTGTAA	779							
QY	783	GATCAGTAATACGCGCAACTGTTGAATGTGAATGTTTGAATACTGGAAGTGAAGC	842							
DB	780	GATCAGTAATACGCGCAACTGTTGAATGTGAATGTTTGAATACTGGAAGTGAAGC	839							
QY	843	ATGTGACATTCCTCACTGTACAGCAACTGTGGTTTCTCATCGAGGCAATCGCAATTC	902							
DB	840	ATGTGACATTCCTCACTGTACAGCAACTGTGGTTTCTCATCGAGGCAATCGCAATTC	899							
QY	903	AAGTGATGTACAGAGATGCTCCTGCTTCTCAGACTGGCAGGGTCTCGATGTTTCAGTTCC	962							
DB	900	AAGTGATGTACAGAGATGCTCCTGCTTCTCAGACTGGCAGGGTCTCGATGTTTCAGTTCC	959							

QY	963	TGTACCAAGCTAACCAAGTCATTTTGGACTCGAGAGAAATATTTAACTTAAAGCTCCCCAG	1022
DB	960	TGTACCAAGCTAACCAAGTCATTTTGGACTCGAGAGAAATATTTAACTTAAAGCTCCCCAG	1019
QY	1023	AGCATCTCATAAAGCTGTGGTCAATGGAACAATATATGTTGGTGTGTTGGAGATATATGTT	1082
DB	1020	AGCATCTCATAAAGCTGTGGTCAATGGAACAATATATGTTGGTGTGTTGGAGATATATGTT	1079
QY	1083	CAACCACTCAGATTATAACATGTTCTAGCGTATGACCTTCTAGGAGTGGCTTCC	1142
DB	1080	CAACCACTCAGATTATAACATGTTCTAGCGTATGACCTTCTAGGAGTGGCTTCC	1139
QY	1143	ACTAAACCGTCTGTGAAACAATGTTGTTAGATATGTTGATCTTTGGCATATACAA	1202
DB	1140	ACTAAACCGTCTGTGAAACAATGTTGTTAGATATGTTGATCTTTGGCATATACAA	1199
QY	1203	GGATATAAATTTACATGTTATGAGAGAAAATTTGATCCAACTGGGAATGTGACCAATGAGTT	1262
DB	1200	GGATATAAATTTACATGTTATGAGAGAAAATTTGATCCAACTGGGAATGTGACCAATGAGTT	1259
QY	1263	GAGAGTTTTTACATTAATGATGATGTTGTTGACCCCTTAAGCAAAAGGAGCA	1322
DB	1260	GAGAGTTTTTACATTAATGATGATGTTGTTGACCCCTTAAGCAAAAGGAGCA	1319
QY	1323	GTATGTCAGTGGTGGGCACTCTGCACACATTTGTTACACTGAAGAATGGCGAGTGGTCAT	1382
DB	1320	GTATGTCAGTGGTGGGCACTCTGCACACATTTGTTACACTGAAGAATGGCGAGTGGTCAT	1379
QY	1383	GCTGGTCATCTTTGGTCACTCTGCACACATTTGTTACACTGAAGAATGGCGAGTGGTCAT	1442
DB	1380	GCTGGTCATCTTTGGTCACTCTGCACACATTTGTTACACTGAAGAATGGCGAGTGGTCAT	1439
QY	1443	TTTGGATGAACAACACATGAGTATATACACACCCAGGGTGGCTTGTCAAGGGGGTTA	1502
DB	1440	TTTGGATGAACAACACATGAGTATATATACACACCCAGGGTGGCTTGTCAAGGGGGTTA	1499
QY	1503	CGGCCCATAGCAGTGGTTTACGACCATAGGACCGGGCCCTTATACGTTTCACTGGTGGCTACAA	1562
DB	1500	CGGCCCATAGCAGTGGTTTACGACCATAGGACCGGGCCCTTATACGTTTCACTGGTGGCTACAA	1559
QY	1563	GGCTTTTCACTGGCAATAAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGATAC	1622
DB	1560	GGCTTTTCACTGGCAATAAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGATAC	1619
QY	1623	CCAGATGTGGACCATCTTTAAGGACAGCGCGATTTTCCGTTACTTTGCACACAGCTGTGAT	1682
DB	1620	CCAGATGTGGACCATCTTTAAGGACAGCGCGATTTTCCGTTACTTTGCACACAGCTGTGAT	1679
QY	1683	AGTGAGTGAAACCACTGCTGGTGTGTTGGGGGAAACACACAATGACACATCTATGAGCA	1742
DB	1680	AGTGAGTGAAACCACTGCTGGTGTGTTGGGGGAAACACACAATGACACATCTATGAGCA	1739
QY	1743	TGGCGGCAAAATGCTTCTCTTTCAGATTTTATGCGCTTATGACATTTGTCGCTGGTC	1802
DB	1740	TGGCGGCAAAATGCTTCTCTTTCAGATTTTATGCGCTTATGACATTTGTCGCTGGTC	1799
QY	1803	AGTGCTTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTCTT	1862
DB	1800	AGTGCTTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTCTT	1859
QY	1863	ACACACAGCACCATGATGTTGCTGGTGTGTTTCAATAGTCTCTCTCCCTCAGGACATCT	1922
DB	1860	ACACACAGCACCATGATGTTGCTGGTGTGTTTCAATAGTCTCTCTCCCTCAGGACATCT	1919
QY	1923	GGTATTACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTATTAGCAGCAGG	1982
DB	1920	GGTATTACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTATTAGCAGCAGG	1979
QY	1983	ACCTGGTATTGGTGTGTTGGACACAGCGGTGCTCTCAGTGTATCTCGTGGGCGCTGGC	2042
DB	1980	ACCTGGTATTGGTGTGTTGGACACAGCGGTGCTCTCAGTGTATCTCGTGGGCGCTGGC	2039



QY 2043 AACGTGATGAACAGAAAGTTAAATCAGAATGTTTTCCAAAGAACTCTTGACCA 2102  
Db 2040 AACTGATGAACAGAAAGTTAAATCAGAATGTTTTCCAAAGAACTCTTGACCA 2099  
QY 2103 TGACAGATGTGACCAAGCACACAGATTGTTACAGCTGTACAGCAACACCAATGACTGCCA 2162  
Db 2100 TGACAGATGTGACCAAGCACACAGATTGTTACAGCTGTACAGCAACACCAATGACTGCCA 2159  
QY 2163 CTGGTCAATGACCAATGTGCTCCCAAGGACACACAGCTGCTCAGAGGCCAGATCCCAT 2222  
Db 2160 CTGGTCAATGACCAATGTGCTCCCAAGGACACACAGCTGCTCAGAGGCCAGATCCCAT 2219  
QY 2223 TTTTAGGTATGAGAAATGTCGCCCAAGGATAACCCCATGTACTACTGTAAACAAGAACCCAG 2282  
Db 2220 TTTTAGGTATGAGAAATGTCGCCCAAGGATAACCCCATGTACTACTGTAAACAAGAACCCAG 2279  
QY 2283 CTGCAGAGCTGTGCTCCAGGACAGAACTGCCAGTGGGAGCCCGGAAATCAGAGTGCAT 2342  
Db 2280 CTGCAGAGCTGTGCTCCAGGACAGAACTGCCAGTGGGAGCCCGGAAATCAGAGTGCAT 2339  
QY 2343 TGCCCTGCCCGAAAAATATCTGTGGCAATGCTGGCAATTTGGTTGGAAACTCATGTTGAA 2402  
Db 2340 TGCCCTGCCCGAAAAATATCTGTGGCAATGCTGGCAATTTGGTTGGAAACTCATGTTGAA 2399  
QY 2403 AATTACTACTGCCAAGGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAACAATGC 2462  
Db 2400 AATTACTACTGCCAAGGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAACAATGC 2459  
QY 2463 CCTTTGGCTTCTCTTAACAACCCAGAAAGAGTGTGCTTAAAGTGTGCTTAAAGCAGTCCGAAT 2522  
Db 2460 CCTTTGGCTTCTCTTAACAACCCAGAAAGAGTGTGCTTAAAGTGTGCTTAAAGCAGTCCGAAT 2519  
QY 2523 AATGCACTCATCTCAGAGCATGTCCAGCTCACTTAAACCCATGGTGGGCTTCGGAA 2582  
Db 2520 AATGCACTCATCTCAGAGCATGTCCAGCTCACTTAAACCCATGGTGGGCTTCGGAA 2579  
QY 2583 GATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCCAATTTACAAATAGTTTACTACA 2642  
Db 2580 GATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCCAATTTACAAATAGTTTACTACA 2639  
QY 2643 GTGGATGCCCTCTGAGCCAGTGTGCTGATCTGATCTGCTGGAATTTCTGGAATTTATCAGAACCCAGTAC 2702  
Db 2640 GTGGATGCCCTCTGAGCCAGTGTGCTGATCTGATCTGCTGGAATTTCTGGAATTTATCAGAACCCAGTAC 2699  
QY 2703 TCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCC 2762  
Db 2700 TCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCC 2759  
QY 2763 TGCAACCAACAGTCTAAGCAGTCCGACACCATGTGCTTGAAGACAGCATGTGGAGA 2822  
Db 2760 TGCAACCAACAGTCTAAGCAGTCCGACACCATGTGCTTGAAGACAGCATGTGGAGA 2819  
QY 2823 TTGCACCAAGGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGAGACTC 2882  
Db 2820 TTGCACCAAGGGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGAGACTC 2879  
QY 2883 CAATGCCATATGTGGCCCTCTCCCTTTGGCCAGTGTATGGAATGGTATACGATGAGCAC 2942  
Db 2880 CAATGCCATATGTGGCCCTCTCCCTTTGGCCAGTGTATGGAATGGTATACGATGAGCAC 2939  
QY 2943 CTGCCCCCTCGAAATTTGTTGAGGCTACTGTACTGTAGTCAATTTGCTTGGAGCAACCCAGG 3002  
Db 2940 CTGCCCCCTCGAAATTTGTTGAGGCTACTGTACTGTAGTCAATTTGCTTGGAGCAACCCAGG 2999  
QY 3003 CTGTGGCTGTGTACTGATCCAGCAATATCTGGCAAGGGAATGATAGAGGGTTCCTA 3062  
Db 3000 CTGTGGCTGTGTACTGATCCAGCAATATCTGGCAAGGGAATGATAGAGGGTTCCTA 3059  
QY 3063 TAAAGGACCAAGTGAAGTGTCTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCT 3122  
Db 3060 TAAAGGACCAAGTGAAGTGTCTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCT 3119  
QY 3123 GCTCAATTCAGCAATGTGTAGAGGACAGCAGATACAACTGGTCTTTCAATTCATGTGCC 3182

Db 3120 GCTCAATTCAGCAATGTGTAGAGGACAGCAGATACAACTGGTCTTTCAATTCATGTGCC 3179  
QY 3183 AGTTGCCAATGCAACGGCCACAGTAATGCAATGCAATCAGACATCTGTGAGAAATGTGA 3242  
Db 3180 AGTTGCCAATGCAACGGCCACAGTAATGCAATCAGACATCTGTGAGAAATGTGA 3239  
QY 3243 GAACCTTGACCAACAGGCAAGCACTGGAGACCTGCAATCTGGCTTCTACGGTGTATCCAC 3302  
Db 3240 GAACCTTGACCAACAGGCAAGCACTGGAGACCTGCAATCTGGCTTCTACGGTGTATCCAC 3299  
QY 3303 CAATGAGGGAATATGTACGCCATGCAAGTGAATGGGCAAGGCTCTGTGCAACACCAA 3362  
Db 3300 CAATGAGGGAATATGTACGCCATGCAAGTGAATGGGCAAGGCTCTGTGCAACACCAA 3359  
QY 3363 CACGGGCAAGTGTCTTCACCAACACCAAGGCGCTCAAGGGGACAGTGCAGCTATGTGA 3422  
Db 3360 CACGGGCAAGTGTCTTCACCAACCAAGGCGCTCAAGGGGACAGTGCAGCTATGTGA 3419  
QY 3423 GGTAGAAATTCGATACCAAGGAAACCCCTCTCAGAGAAACATGTTATTTACTCTTTTAT 3482  
Db 3420 GGTAGAAATTCGATACCAAGGAAACCCCTCTCAGAGAAACATGTTATTTACTCTTTTAT 3479  
QY 3483 TGACTATCAGTTTCACTTGTAGTCTATCCAGGAAGATGATCGCTATTCACAGCTATCAA 3542  
Db 3480 TGACTATCAGTTTCACTTGTAGTCTATCCAGGAAGATGATCGCTATTCACAGCTATCAA 3539  
QY 3543 TTTTGTGGCTTACTCTGACGAAACAAACAGGGAATTTGGACATGTTCAATATGCCCTCAA 3602  
Db 3540 TTTTGTGGCTTACTCTGACGAAACAAACAGGGAATTTGGACATGTTCAATATGCCCTCAA 3599  
QY 3603 GAATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3662  
Db 3600 GAATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3659  
QY 3663 AGAGATGCTCTGTGTTTCAAAAAACCAATTAAGGAGTACAAAGATAGTTTCTCTAAATGA 3722  
Db 3660 AGAGATGCTCTGTGTTTCAAAAAACCAATTAAGGAGTACAAAGATAGTTTCTCTAAATGA 3719  
QY 3723 GAAGTTTGAATTTTGGCAACCCCAATATCACTTTCTTTGTTATGTGCAATTTTCAAC 3782  
Db 3720 GAAGTTTGAATTTTGGCAACCCCAATATCACTTTCTTTGTTATGTGCAATTTTCAAC 3779  
QY 3783 CTGGCCCATCAAAATTCAGATTGCCCTCTCTCAGCAACAGCAATTTTATGGACCTGGTACA 3842  
Db 3780 CTGGCCCATCAAAATTCAGATTGCCCTCTCTCAGCAACAGCAATTTTATGGACCTGGTACA 3839  
QY 3843 GTTCTTCGTGACTTCTCAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3902  
Db 3840 GTTCTTCGTGACTTCTCAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3899  
QY 3903 GAAGATCAAAACAAAGTTGTTGGGCTTCCAGAGTGAAGCAACTTCTTCGAGAGATGCA 3962  
Db 3900 GAAGATCAAAACAAAGTTGTTGGGCTTCCAGAGTGAAGCAACTTCTTCGAGAGATGCA 3959  
QY 3963 ACAGATGGCCAGCGCTCTTTCGCTCTGTAATGTCGCTTGGAAACAGATGAGAGGCC 4022  
Db 3960 ACAGATGGCCAGCGCTCTTTCGCTCTGTAATGTCGCTTGGAAACAGATGAGAGGCC 4019  
QY 4023 TCCTGATCTTATTTGGGGAGTATAAGACTCTTCCAAACCCATTTGCACTCGAGCCCGTG 4082  
Db 4020 TCCTGATCTTATTTGGGGAGTATAAGACTCTTCCAAACCCATTTGCACTCGAGCCCGTG 4079  
QY 4083 TTTTGGCAACAGCGCT 4142  
Db 4080 TTTTGGCAACAGCGCT 4139  
QY 4143 CATCCCT 4202  
Db 4140 CATCCCT 4199  
QY 4203 GCAGATGCCAGTGTCTCAAGGAGAGTCAAGGAGCCGTGAGAAACCCGAGACAGCAGCC 4262

Db	4200	GCAGATCGGATAGTGTACAGGAGAGAGTCCAGGAGCGGTGAGAAACCGGAGCAGAGCC	4259
Qy	4263	CCCTGACAGCCTGGGACCTGCTCTGA	4290
Db	4260	CCCTGACAGCCTGGGACCTGCTCTGA	4287
RESULT 2			
AF106861			
LOCUS	AF106861	3819 bp	mRNA linear PRI 06-JAN-1999
DEFINITION	Homo sapiens attractin-2 (ATRN) mRNA, complete cds.		
ACCESSION	AF106861		
VERSION	AF106861.1		GI:4093195
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 3819)		
TITLE	Duke-Cohan, J.S., Gu, J., Freeman, G.J. and Schlossman, S.F.		
JOURNAL	Cloning of cDNA for attractin-2, identical with that of attractin except for a GC-rich 222 bp 5' insertion		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 3819)		
TITLE	Duke-Cohan, J.S., Gu, J., Freeman, G.J. and Schlossman, S.F.		
JOURNAL	Direct Submission		
FEATURES	Submitted (15-NOV-1998) Division of Tumor Immunology, Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA		
source	Location/Qualifiers		
gene	1..3819		
CDS	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	1..3819		
	/gene="ATRN"		
	/gene="ATRN"		
	/note="similar to Homo sapiens attractin in GenBank"		
	Accession Number AF034957"		
	/codon_start=1		
	/product="attractin-2"		
	/protein_id="AAD03057.1"		
	/db_xref="GI:4093196"		
	/translation="MVAATAATEARLRRTTAAALAGRCGGHMDWDVTRAGPGLG AGLRPLRLSPPLRLDLLLPPPLLLLECEABAAAAAASGSAABAEKCD RFLVNGGRCNPGTCQCPAGVGEQCHQCGRPRLTGSQFVTDGPNYKTKTW LIEGPNRIMLRNFHAFETCSWDHLYVVDGSIYAPLVAASGLIIPERDNETVPE VWATSGYALLHFFSDAAVNLGENTYISFDMPCNNCSRGCECKISNSSETVECESEN WKGACDIPHCDCNCGPHEGICNSDVRGCSFSDQPGGCVFPVFNQSFWRREY SNLKLPRASHAVNGNIMWVVGTFNFDNDYNNVLAIDLASREWLPLNRSVNVVVR YGHSLATKIKIYMGKIDPTGNVTELRVFIHNSVLLTPKAKEQYAVVGHSAH LVLKNGVVMVLFHGCPLXYISNVQEDLDKNTWSILHTQCALVQGGYHSSVVD HRTALYVHGKYAFSAKRYLADLYVDVDTQMTILKDSRFFRYLHTAVTVSGTM LVFGQTHNDTSMHGAKCFSSDPWYDIACDRWSVLPRLDHDVNRFGHSVLGNS TMVYFGGNSLLSDILVFTSEQDAHRSAAALAGPGRICVWNTGSSQICSWALAT DEOREKSECFKRLDHDRCDDTDCYSTANTNDCHNCNHCUPRNHSCSGQLS LFRVENCPCPNXYCNKTKSCRSALDQNCWEPNQCIALPENICIGIGHLVNS CLKITAKENYDNLAKLFRHNHALLSLTKQKVEFVLKQRLRMQSSQSKLLTPW GSLRKINVSWMEDMFPFNTSLQWMPSEPSDAGFCGILSEPTSLKLAATCINPLN GVCERPANHSKACRTPCALRTACGDCSTSGSECMCMNKCVDNSNAYVAFSPFQ CMETWYMTCTPPENCSCYCTCHCLEQPGGWCCTDPNTGKGCISGYSKGPVPMQSO APTGNFVPOPLLNSMCLDSRYNWSFIHCPACQNGSKCINQSIKCEKNLJTKKH CETCISGFGDPTNGKCPCKNGKSHASLNTWTKFCFTKGVKDEBQCLCHEVRY QGNPLRCTCYLLIDYQTFSLSQEDDRYTAINFVATPEQNRDLDFINASKNFN LNIWASFSAGTQAGEMPVWSKNIKEYKDSFNEKDFRNPHTFFVYVNSFTW PIKIQVOTEQ"		
ORIGIN			
Query Match	88.6%; Score 3801.4; DB 9; Length 3819;		
Best Local Similarity	100.0%; Pred. No. 0;		
Matches 3802; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
Qy	1	ATGTGGCCGCGCGGCAACTGAGGCAAGCTGAGGAGGAGGAGCGCGGCGACGCA	60

Db 1081 TTCAACCACTCAGATTATTAACATGGTTCTAGCGTATGACCTTGTCTTAGGAGTGGCTT 1140  
Qy 1141 CCACTAAACCGTCTCTGTGAACAATGTGGTTGTAGATATGGTCATTTTGGCATATAC 1200  
Db 1141 CCACTAAACCGTCTCTGTGAACAATGTGGTTGTAGATATGGTCATTTTGGCATATAC 1200  
Qy 1201 AAGGATAAATTTACATGTATGGAGGAAAAATGTATCCAACTGGGAATGTGACCAATGAG 1260  
Db 1201 AAGGATAAATTTACATGTATGGAGGAAAAATGTATCCAACTGGGAATGTGACCAATGAG 1260  
Qy 1261 TTGAGAGTGTTCACATTTTACATATGAGTCATGGGTGTCTGTGACCCCTTAAGGCAAAAGGAG 1320  
Db 1261 TTGAGAGTGTTCACATTTTACATATGAGTCATGGGTGTCTGTGACCCCTTAAGGCAAAAGGAG 1320  
Qy 1321 CAGTATGCAATGTGGGCACTCTGCACACATTTGTACCTGAAAGATGGCCGAGTGTCT 1380  
Db 1321 CAGTATGCAATGTGGGCACTCTGCACACATTTGTACCTGAAAGATGGCCGAGTGTCT 1380  
Qy 1381 ATGCTGTCTCATTTTGGTCACTGCCCCCTCTATGGATATATAAGCAATGTGACGAATAT 1440  
Db 1381 ATGCTGTCTCATTTTGGTCACTGCCCCCTCTATGGATATATAAGCAATGTGACGAATAT 1440  
Qy 1441 GATTTGGATAGAACACATGGAGTATATTACACACCCAGGTTGCCCTTGTGCAAGGGGT 1500  
Db 1441 GATTTGGATAGAACACATGGAGTATATTACACACCCAGGTTGCCCTTGTGCAAGGGGT 1500  
Qy 1501 TAGCGGCATAGCAGTGTTTACGACCATAGGACCCAGGGCCCTATACGTTTCATGGTGGCTAC 1560  
Db 1501 TAGCGGCATAGCAGTGTTTACGACCATAGGACCCAGGGCCCTATACGTTTCATGGTGGCTAC 1560  
Qy 1561 AAGGCTTTAGTGCCCAATAGATACCGGCTTGCGATGATCTCTACGGATATGATGTGAT 1620  
Db 1561 AAGGCTTTAGTGCCCAATAGATACCGGCTTGCGATGATCTCTACGGATATGATGTGAT 1620  
Qy 1621 ACCCAGATGGGACCACTTTTAAGGACAGCCGATTTTCCGTTACTTTCGCTTACTTGACACAGCTGTG 1680  
Db 1621 ACCCAGATGGGACCACTTTTAAGGACAGCCGATTTTCCGTTACTTTCGCTTACTTGACACAGCTGTG 1680  
Qy 1681 ATAGTGTGGAACCAATGTGGTGTGGGGGAACACACACAAATGACACATCTATGAGC 1740  
Db 1681 ATAGTGTGGAACCAATGTGGTGTGGGGGAACACACACAAATGACACATCTATGAGC 1740  
Qy 1741 CATGGGCGCAAAATGCTTCTTCAGATTTTCATGGCCATATGACATTTGCCCTGTGACCGCTGG 1800  
Db 1741 CATGGGCGCAAAATGCTTCTTCAGATTTTCATGGCCATATGACATTTGCCCTGTGACCGCTGG 1800  
Qy 1801 TCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860  
Db 1801 TCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860  
Qy 1861 TTACACAAAGCAGCAGTATGTGTGGTGGTTTCAATAGTCTCTCTCTCAGCGACATC 1920  
Db 1861 TTACACAAAGCAGCAGTATGTGTGGTGGTTTCAATAGTCTCTCTCTCAGCGACATC 1920  
Qy 1921 CTGGTATTCACCTCGGAACAGTGTGATGGGCAATCGGAGTGAAGCGCTTGTATACAGCA 1980  
Db 1921 CTGGTATTCACCTCGGAACAGTGTGATGGGCAATCGGAGTGAAGCGCTTGTATACAGCA 1980  
Qy 1981 GGACCTGGTATTCGGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCTGGGCGGTG 2040  
Db 1981 GGACCTGGTATTCGGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCTGGGCGGTG 2040  
Qy 2041 GCAACTGATGAAACAAGAAAAAGTTAAATCAGATTTTTCAAAAGAACTCTTGAC 2100  
Db 2041 GCAACTGATGAAACAAGAAAAAGTTAAATCAGATTTTTCAAAAGAACTCTTGAC 2100  
Qy 2101 CATGACAGATGTGACACAGCAGACAGATTTGTTACAGCTGTACAGCAACCAATCAGCTGC 2160  
Db 2101 CATGACAGATGTGACACAGCAGACAGATTTGTTACAGCTGTACAGCAACCAATCAGCTGC 2160  
Qy 2161 CACTGTGCAATGACCAATTTGTGCCCAAGAACACAGCTGTCTCAGAGGCGCAGATCTCC 2220  
Db 2161 CACTGTGCAATGACCAATTTGTGCCCAAGAACACAGCTGTCTCAGAGGCGCAGATCTCC 2220

Qy 2221 ATTTTATGATGAGAAATGCCCCAAGGATACCCCCATGATCTACTGTAAACAAGAGACC 2280  
Db 2221 ATTTTATGATGAGAAATGCCCCAAGGATACCCCCATGATCTACTGTAAACAAGAGACC 2280  
Qy 2281 AGCTGCAGAGCTGTGCCCTGACCAAGCACTGCGAGTGGGAGCCCGGAATCAGAGTGC 2340  
Db 2281 AGCTGCAGAGCTGTGCCCTGACCAAGCACTGCGAGTGGGAGCCCGGAATCAGAGTGC 2340  
Qy 2341 ATTGCCCTCCCGAAAAATATCTGTGCAATGCGCTGGCAATTTGGTTGGAAATCTCATGTTG 2400  
Db 2341 ATTGCCCTCCCGAAAAATATCTGTGCAATGCGCTGGCAATTTGGTTGGAAATCTCATGTTG 2400  
Qy 2401 ABAATTTACTACTGCCAAGGAGAAATATGACANTGCTAAATTTGTTGTAGGAACCAAT 2460  
Db 2401 ABAATTTACTACTGCCAAGGAGAAATATGACANTGCTAAATTTGTTGTAGGAACCAAT 2460  
Qy 2461 GCCCTTTTGGCTTCTCTTAAACCCAGAAAGTAGAATTTGTCTTTAAGCAGCTGCGA 2520  
Db 2461 GCCCTTTTGGCTTCTCTTAAACCCAGAAAGTAGAATTTGTCTTTAAGCAGCTGCGA 2520  
Qy 2521 ATAAATCAGTCACTCAGAGCATGTCAGCTCACCTTAAACCCATGGTGGCTTCGG 2580  
Db 2521 ATAAATCAGTCACTCAGAGCATGTCAGCTCACCTTAAACCCATGGTGGCTTCGG 2580  
Qy 2581 AAGATCAATGTGTCTTACTGCTGGGAAGATATGTCCTTAAACCAATAGTTTACTA 2640  
Db 2581 AAGATCAATGTGTCTTACTGCTGGGAAGATATGTCCTTAAACCAATAGTTTACTA 2640  
Qy 2641 CAGTGAATGCCCTGAGCCGATGATGCTGANTCTGGAATTTTATCAGAACCCAGT 2700  
Db 2641 CAGTGAATGCCCTGAGCCGATGATGCTGANTCTGGAATTTTATCAGAACCCAGT 2700  
Qy 2701 ACTCGGGGACTGAAGGCTGCAATCAACCCACTCAATGGTAGTGTCTGTGAAAG 2760  
Db 2701 ACTCGGGGACTGAAGGCTGCAATCAACCCACTCAATGGTAGTGTCTGTGAAAG 2760  
Qy 2761 CTTGCAAAACAAGTCTTAAGCAGTCCCGACACCATGCTGCTGAGCAGCATGTGGA 2820  
Db 2761 CTTGCAAAACAAGTCTTAAGCAGTCCCGACACCATGCTGCTGAGCAGCATGTGGA 2820  
Qy 2821 GATTGACACAGGGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGAC 2880  
Db 2821 GATTGACACAGGGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGAC 2880  
Qy 2881 TCCAATGCTATGTGGCTCCCTTCCCTTTGGCCAGTGTATGGAATGATACAGATGAC 2940  
Db 2881 TCCAATGCTATGTGGCTCCCTTCCCTTTGGCCAGTGTATGGAATGATACAGATGAC 2940  
Qy 2941 ACCTGCCCCCTGAAAAATTTTTCAGGCTACTGCTAGTCAATTTGCTTGGAGCAACA 3000  
Db 2941 ACCTGCCCCCTGAAAAATTTTTCAGGCTACTGCTAGTCAATTTGCTTGGAGCAACA 3000  
Qy 3001 GGCTGTGGCTGTGATGATCCAGCAATACTGCAAGGGAATGATAGAGGTTCC 3060  
Db 3001 GGCTGTGGCTGTGATGATCCAGCAATACTGCAAGGGAATGATAGAGGTTCC 3060  
Qy 3061 TATAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAAGGCC 3120  
Db 3061 TATAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAAGGCC 3120  
Qy 3121 CTGCTCAATTCAGATGTGTCTAGAGNACAGCATCACTGTCTTTTCACTCAGTGT 3180  
Db 3121 CTGCTCAATTCAGATGTGTCTAGAGNACAGCATCACTGTCTTTTCACTCAGTGT 3180  
Qy 3181 CCAGCTTGCAATGCAACCGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGT 3240  
Db 3181 CCAGCTTGCAATGCAACCGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGT 3240  
Qy 3241 GAGAACTGACACAGGGAAGCAGCTGCGAGACCTGATATCTGGCTTCTACGGTATCCC 3300  
Db 3241 GAGAACTGACACAGGGAAGCAGCTGCGAGACCTGATATCTGGCTTCTACGGTATCCC 3300



QY 1263 GAGAGTTTTTACATTCATATGAGTCATGGGTGTTGTTGACCCCTTAAGCCAAAGAGCA 1322  
Db 1260 GAGAGTTTTTACATTCATATGAGTCATGGGTGTTGTTGACCCCTTAAGCCAAAGAGCA 1319  
QY 1323 GTATGCAAGTGTGGGCACTCTGCACACATTTGTTACATGAAGATGGCGAGTGTCTAT 1382  
Db 1320 GTATGCAAGTGTGGGCACTCTGCACACATTTGTTACATGAAGATGGCGAGTGTCTAT 1379  
QY 1383 GCTGGTCATCTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGA 1442  
Db 1380 GCTGGTCATCTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGA 1439  
QY 1443 TTTGGAATAAGAACACATGAGTATATATACACACCCAGGGTCCCTTGTCAAGGGGGTTA 1502  
Db 1440 TTTGGAATAAGAACACATGAGTATATATACACACCCAGGGTCCCTTGTCAAGGGGGTTA 1499  
QY 1503 CGGCCATAGCAGTGTATACACCATAGGACCAAGGCCCTTATACGTTTATGGTGGCTACAA 1562  
Db 1500 CGGCCATAGCAGTGTATACACCATAGGACCAAGGCCCTTATACGTTTATGGTGGCTACAA 1559  
QY 1563 GCTTTTCAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGGATAC 1622  
Db 1560 GCTTTTCAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGGATAC 1619  
QY 1623 CCAGATGTGACCATCTTTAAGGACACCGCATTTTTCCGTTACTTGCACACAGCTGTGAT 1682  
Db 1620 CCAGATGTGACCATCTTTAAGGACACCGCATTTTTCCGTTACTTGCACACAGCTGTGAT 1679  
QY 1683 AGTCAGTGGACCAATGCTGTGTGTTGGGGAAACACACACATGACACATCTATGAGCA 1742  
Db 1680 AGTCAGTGGACCAATGCTGTGTGTTGGGGAAACACACACATGACACATCTATGAGCA 1739  
QY 1743 TGGCGCCAAATGCTTCTTTTTCAGATTTTCAATGGCTATGACATTTGCTGTGACCGTGGTC 1802  
Db 1740 TGGCGCCAAATGCTTCTTTTTCAGATTTTCAATGGCTATGACATTTGCTGTGACCGTGGTC 1799  
QY 1803 AGTCCTTCCAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGAGCTCTT 1862  
Db 1800 AGTCCTTCCAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGAGCTCTT 1859  
QY 1863 ACACACAGACCATGATGTGTGTTGGGTGTTTCAATAGTCTCTCTCCAGCGACATCTT 1922  
Db 1860 ACACACAGACCATGATGTGTGTTGGGTGTTTCAATAGTCTCTCTCCAGCGACATCTT 1919  
QY 1923 GGTATTCACCTCGGAACAGTGTGATGGCATCGGATGAGCGCTGTTTACGACGAGG 1982  
Db 1920 GGTATTCACCTCGGAACAGTGTGATGGCATCGGATGAGCGCTGTTTACGACGAGG 1979  
QY 1983 ACCTGGTATTCGGTGTGTGGAAACAAGGTCGTCTCAGTGTATCTCGTGGCGCTGCG 2042  
Db 1980 ACCTGGTATTCGGTGTGTGGAAACAAGGTCGTCTCAGTGTATCTCGTGGCGCTGCG 2039  
QY 2043 AACTGATGAACAAGAAAGTTAAATCAGATGTTTTCAAAAGAACTCTTGACCA 2102  
Db 2040 AACTGATGAACAAGAAAGTTAAATCAGATGTTTTCAAAAGAACTCTTGACCA 2099  
QY 2103 TGACAGATGTGACCAAGACACAGATTTTACAGCTGTACAGCAACACCAATGATGACCA 2162  
Db 2100 TGACAGATGTGACCAAGACACAGATTTTACAGCTGTACAGCAACACCAATGATGACCA 2159  
QY 2163 CTGGTGAATGACCAATTTGTCCTCCAGGAACCAAGCTGCTCAGAGGCCAGATCTCCAT 2222  
Db 2160 CTGGTGAATGACCAATTTGTCCTCCAGGAACCAAGCTGCTCAGAGGCCAGATCTCCAT 2219  
QY 2223 TTTTATGATGAGAAATGGCCCCAAGGATAACCCCATGTTACTCTGTAAACAAGAACCCAG 2282  
Db 2220 TTTTATGATGAGAAATGGCCCCAAGGATAACCCCATGTTACTCTGTAAACAAGAACCCAG 2279  
QY 2283 CTGACAGAGCTGTGCTTGCACCAAGAACTGCAAGTGGAGCCCGGAGTATGAGTGTGAT 2342  
Db 2280 CTGACAGAGCTGTGCTTGCACCAAGAACTGCAAGTGGAGCCCGGAGTATGAGTGTGAT 2339  
QY 2343 TGCCCTGCCGAAAAATATCTGTGGCATTTGGTGGCAATTTGGTGGAAACTCATGTTTGA 2402

Db 2340 TGCCCTGCCGAAAAATATCTGTGGCATTTGGTGGCAATTTGGTGGAAACTCATGTTTGA 2399  
QY 2403 AATTACTACTGCCAAGAGAAATTTATGACAAATGCTAAATTTGTTCTGTAGGAACCAAAATGC 2462  
Db 2400 AATTACTACTGCCAAGAGAAATTTATGACAAATGCTAAATTTGTTCTGTAGGAACCAAAATGC 2459  
QY 2463 CTTTTGGCTTCTCTTACAAACCCAGAAAGGTAGAAATTTGCTTTAAGCAGCTGCGAAT 2522  
Db 2460 CTTTTGGCTTCTCTTACAAACCCAGAAAGGTAGAAATTTGCTTTAAGCAGCTGCGAAT 2519  
QY 2523 AATGCACTCATCTCAGAGCATGTCCAAAGCTCACTTTAACCCCATGGTGGCTTCGGAA 2582  
Db 2520 AATGCACTCATCTCAGAGCATGTCCAAAGCTCACTTTAACCCCATGGTGGCTTCGGAA 2579  
QY 2583 GATCAATGTGCTCTAATGCTGTGGGAAGATATGCTGCCCATTTAACAATAGTTTACTACA 2642  
Db 2580 GATCAATGTGCTCTAATGCTGTGGGAAGATATGCTGCCCATTTAACAATAGTTTACTACA 2639  
QY 2643 GTGATGSCCTCTGAGCCAGTGTCTGATCTGATTTCTGTGGAAATTTTATCAGAACCCAGTAC 2702  
Db 2640 GTGATGSCCTCTGAGCCAGTGTCTGATTTCTGTGGAAATTTTATCAGAACCCAGTAC 2699  
QY 2703 TCGGGGAATGAAGGCTGGAACCTGTGCATCAACCCCATCAATGTTAGTGTCTGTGAAGGCC 2762  
Db 2700 TCGGGGAATGAAGGCTGGAACCTGTGCATCAACCCCATCAATGTTAGTGTCTGTGAAGGCC 2759  
QY 2763 TGCAAAACCAAGTGTAAAGCAGTGGCGGACACCATGTGCTTTGAGCAGCATGTGGAGA 2822  
Db 2760 TGCAAAACCAAGTGTAAAGCAGTGGCGGACACCATGTGCTTTGAGCAGCATGTGGAGA 2819  
QY 2823 TTGCAACAGCGGAGCTGTGAGTGTGTCAGCAACATGAAGCAAGTGTGTGGACTC 2882  
Db 2820 TTGCAACAGCGGAGCTGTGAGTGTGTCAGCAACATGAAGCAAGTGTGTGGACTC 2879  
QY 2883 CAATGCTATGTGCTCTTCCCTTTGSCAGTGTATGGAATGATATACATGAGCAGC 2942  
Db 2880 CAATGCTATGTGCTCTTCCCTTTGSCAGTGTATGGAATGATATACATGAGCAGC 2939  
QY 2943 CTGCCCCCTGAAAAATGTTTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCCAGG 3002  
Db 2940 CTGCCCCCTGAAAAATGTTTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCCAGG 2999  
QY 3003 CTGTGCTGTGTACTGATCCAGCAATACTCGCAAGGGAATGATCATAGAGGTTCTTA 3062  
Db 3000 CTGTGCTGTGTACTGATCCAGCAATACTCGCAAGGGAATGATCATAGAGGTTCTTA 3059  
QY 3063 TAAAGGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3119  
Db 3060 TAAAGGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3119  
QY 3123 GCTCAATTCAGCATGTGTAGAGCAGCAGATACACTGCTTCTTTCATTCAGTGTCC 3182  
Db 3120 GCTCAATTCAGCATGTGTAGAGCAGCAGATACACTGCTTCTTTCATTCAGTGTCC 3179  
QY 3183 AGCTTGCATATGCAAGCGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTGA 3242  
Db 3180 AGCTTGCATATGCAAGCGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTGA 3239  
QY 3243 GAACCTTGACCAAGCAAGCACTGTGAGCAGCTGCAATNTCTGCTTCTAGCGTGTATCCAC 3302  
Db 3240 GAACCTTGACCAAGCAAGCACTGTGAGCAGCTGCAATNTCTGCTTCTAGCGTGTATCCAC 3299  
QY 3303 CAATGGAGGAAATGTGAGCAGTGAAGTGAATGGGCAAGCGCTCTGTGTCAACACCAA 3362  
Db 3300 CAATGGAGGAAATGTGAGCAGTGAAGTGAATGGGCAAGCGCTCTGTGTCAACACCAA 3359  
QY 3363 CAGGGCAGTGTCTTGTGACCAACCAAGCGCTGAGGGGAGCGAGTGCAGCTATGTGA 3422  
Db 3360 CAGGGCAGTGTCTTGTGACCAACCAAGCGCTGAGGGGAGCGAGTGCAGCTATGTGA 3419  
QY 3423 GGTAGAAAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATTACTCTTCTTAT 3482





Qy 913 AGAGATGCTCCCTGCTTCTCAGACTGCGAGGCTCTGGATGTCAGTTCCTGTCACAGCT 972  
Db 925 AGAGATGCTCCCTGCTTCTCAGAGTGGCAAGGTCCTGGATGTCAGTTCCTGTCACAGCT 984  
Qy 973 AACAGTCAATTTGGGACTCGAGAGGAATATCTTAACCTTAAGCTCCCGAGGACCTCAT 1032  
Db 985 AACAGTCAATTTGGACACAGAGGAATATCTTAAGCTTAAGCTCCCGAGGACCTCTCAC 1044  
Qy 1033 AAAGCTGTGTCAATGGAACCAATTAATGTGGTGTGGAGGATATATGTTCAACCACTCA 1092  
Db 1045 AAAGCTGTGTCAATGGAACCAATTAATGTGGTGTGGAGGATATATGTTCAACCACTCA 1104  
Qy 1093 GATTATAACATGCTCTAGGATGACCTGCTTCTAGGAGTGGCTTCCACTAAACCGT 1152  
Db 1105 GATTATAACATGCTCTAGGATGACCTGCTTCTAGGAGTGGCTTCCACTAAACCGT 1164  
Qy 1153 TCTGTGAACAAATGCTGTGTAGATATGTCATCTTTTGGCAATATACAAGGATAAAAT 1212  
Db 1165 TCTGTGAACAAATGCTGTGTAGATATGTCATCTTTTGGCAATATACAAGGATAAAAT 1224  
Qy 1213 TACATGTATGGGAAATATGATCAACTGGGATGTGACCAATGATGAGTGGAGTTTT 1272  
Db 1225 TACATGTATGGGAAATATGATCAACAGGAATGTGACCAATGATGAGTGGAGTTTT 1284  
Qy 1273 CACATTCATAATGAGTCATGGTGTGTTGACCCCTTAAGGCAAAAGGAGCAGTATGCACTG 1332  
Db 1285 CACATTCATAATGAGTCATGGTGTGTTGACCCCTTAAGGCAAAAGGAGCAGTATGCACTG 1344  
Qy 1333 GTTGGGCACTCTGCACACATGTTACACTGAAGATGGCCGAGTGGTCATGCTGGTCATC 1392  
Db 1345 GTTGGGCACTCTGCACACATGTTACACTGAAGATGGCCGAGTGGTCATGCTGGTCATC 1404  
Qy 1393 TTTGGTCACTGCTCTCTATGATATATAGCAATGTGCAAGGAATATGATTTGGATAAG 1452  
Db 1405 TTTGGTCACTGCTCTCTATGATATATAGCAAGTGTGCAAGGAATATGATTTGGATAAG 1464  
Qy 1453 AACACATGGATATATACACACCGAGTGGCTTGTGCAAGGGGTTACGGCCATAGC 1512  
Db 1465 AACACATGGATATATACACACTAGGGTGGCTTGTGCAAGGGGTTACGGCCATAGC 1524  
Qy 1513 AGTGTTCATGATAGGACCAAGGCTGTATATTCAGGAGGCTACAGGCTTTCACT 1572  
Db 1525 AGTGTTCATGATAGGACCAAGGCTGTATATTCAGGAGGCTACAGGCTTTCACT 1584  
Qy 1573 GCCAATAGTACCGGCTTCAGATGATCTCTACCGATATGATGATGATACCGAGATGG 1632  
Db 1585 GCCAATAGTACCGGCTTCAGATGATCTCTACAGATATGATGATGATGATGATGATG 1644  
Qy 1633 ACCATTCCTTAAGACACAGCGATTTTCCGTTACTTGCACACAGCTGTGATAGTGGGA 1692  
Db 1645 ACCATTCCTTAAGACACAGCGATTTTCCGTTACTTGCACACAGCTGTGATAGTGGGA 1704  
Qy 1693 ACCATGCTGTGTTTGGGGAACACACACATGACATCTATGAGGCTATGGCCGCAAA 1752  
Db 1705 ACCATGCTGTGTTTGGGGAACACACACATGACATCTATGAGGCTATGGCCGCAAA 1764  
Qy 1753 TGCTTCTCTTCAGATTCATGGCTATGACATGCTGTGACCGGCTGTGAGTGTCTTCCC 1812  
Db 1765 TGCTTCTCTTCAGATTCATGGCTATGACATGCTGTGACCGGCTGTGAGTGTCTTCCC 1824  
Qy 1813 AGACCTGATCTCCACATGATCAACAGATTTGGCCATTCAGCAGTCTTACACACAGC 1872  
Db 1825 AGACCTGATCTCCACATGATCAACAGATTTGGCCATTCAGCAGTCTTACACACAGC 1884  
Qy 1873 ACCATGATGTGTTCCGTTGGTTTCAATAGTCTTCCCTCCTCAGGACATCTCGTATTCAC 1932  
Db 1885 ACCATGATGTGTTTCCGTTGGTTTCAATAGTCTTCCCTCCTCAGGACATCTCGTATTCAC 1944  
Qy 1933 TCGGAACAGTGTGATCGGATCGAGTGAAGCGGCTTTTGAAGAGGAGCTTCTATTAAGGACCA 1992  
Db 1945 TCGGAACAGTGTGATCGGATCGAGTGAAGCGGCTTTTGAAGAGGAGCTTCTATTAAGGACCA 2004

Qy 1993 CGGTGTGTGGGAACACAGGTCGTCTCAGTGTATCTCTGGGCGCTGGCAACTGATGAA 2052  
Db 2005 CGGTGTGTGTGGGACACAGGTCATCTCAGTGTCTCTCTGGGAGCTGGCACTGAACA 2064  
Qy 2053 CAAGAGAAAGATTAATAATCAGAAATGTTTCCAAAAGAACTCTTGGACCATGACAGATGT 2112  
Db 2065 C---AAGAAAGATTAAGTCAGAAATGTTTCTAAAGAAATTTTGAACATGACAGATGT 2121  
Qy 2113 GACCAGCACACAGATTTGTTTACAGCTGTACAGCAACACCAATGACTGCCACTGGTGCAAT 2172  
Db 2122 GACCAGTCACGATTTGTTTACAGCTGTACAGCAACACCAATGACTGCCACTGGTGCAAT 2181  
Qy 2173 GACCAATGTCCTCCAGGAACACACAGCTGTCTCAGAGGCCAGATCTCCATTTTGGTAT 2232  
Db 2182 GACCAATGTCCTCCAGGAACACACAGCTGTCTCAGAGGCCAGATCTCCATTTTGGTAT 2241  
Qy 2233 GAGAATGTCCTCCAGGAATAACCCATGTACTACTGTAAACAAGAACACAGCTGCGAGGAC 2292  
Db 2242 GATCATTTGTCCTCCAGGAATAACCCATGTATTATTTAAACAAGAACACAGCTGCGAGGAC 2301  
Qy 2293 TGTGCTCTCGACCAACACTGCGAGTGGGAGCCCGGATCAGGATGCAATTCCTCTGCC 2352  
Db 2302 TGTGCTCTCGACCAACACTGCGAGTGGGAGCCCGGATCAGGATGCAATTCCTCTGCC 2361  
Qy 2353 GAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAATCTCATGTTTGAATAATTACTACT 2412  
Db 2362 GAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAATCTCATGTTTGAATAATTACTACT 2421  
Qy 2413 GCCAAGGAAATATGACAACTGTAATTTGTCTGTAGGAACCAATTCCTCTTGGCT 2472  
Db 2422 GCCAAGGAAATATGACAACTGTAATTTGTCTGTAGGAACCAATTCCTCTTGGCT 2481  
Qy 2473 TCTCTTAAACCCAGGAAGGATAGAAATTTGCTTAAAGAGCTGCGAATTAATGAGTCA 2532  
Db 2482 TCTCTTAAACCCAGGAAGGATAGAAATTTGCTTAAAGAGCTGCGAATTAATGAGTCA 2541  
Qy 2533 TCTCAGAGCATCTCCAGCTCACCTTACCCCATGGTGGCTTCCGAAGATCAATGTG 2592  
Db 2542 TACAGGATGTCCTCAAGTTTCACTTAAACCCCATGGTGGCTTCCGAAGATCAATGTG 2601  
Qy 2593 TCTACTGTGTGGGAAGATATGTCCCATTTTCAAAATAGTTTACTACAGTGGATGCG 2652  
Db 2602 TCTACTGTGTGGGAAGATATGTCCCATTTTCAAAATAGTTTACTACAGTGGATGCG 2661  
Qy 2653 TCTGAGCCAGTGTGGATCTGTGAATTTTATCAGAACCCAGTACTCTGGGAGCTG 2712  
Db 2662 TCTGAGCCAGTGTGGATCTGTGAATTTTATCAGAACCCAGTACTCTGGGAGCTG 2721  
Qy 2713 AAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGTGAAAGGCTTGCACCAAC 2772  
Db 2722 AAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGTGAAAGGCTTGCACCAAC 2781  
Qy 2773 AGTGTAGCAGTGGGACACCATGTGCTTGGAGCAGCATGTGGAGATGACACGAC 2832  
Db 2782 AGTGTAGCAGTGGGACACCATGTGCTTGGAGCAGCATGTGGAGATGACACGAC 2841  
Qy 2833 GGCAGCTCTGAGTGTGGTGGCAACATGAAGCAGTGTGGAGTCTCCAAATGCTCTAT 2892  
Db 2842 GGCAGCTCTGAGTGTGGTGGCAACATGAAGCAGTGTGGAGTCTCCAAATGCTCTAT 2901  
Qy 2893 GTGGCTCTCTCTCTTTGGCCAGTGTATGAAATGATATAGATGAGCAGCTGCCCCCT 2952  
Db 2902 GTGGCTCTCTCTCTTTGGCCAGTGTATGAAATGATATAGATGAGCAGCTGCCCCCT 2961  
Qy 2953 GAAAATGTTTCAAGGCTACTGTACTGTAGTCAATTTGCTTGGAGCAACACAGGCTGTGGTGG 3012  
Db 2962 GAAAATGTTTCAAGGCTACTGTACTGTAGTCAATTTGCTTGGAGCAACACAGGCTGTGGTGG 3021  
Qy 3013 TGTACTGATCCAGCAATCTGCGCAAGGAAATGATAGAGGTTCTATTAAGGACCA 3072  
Db 3022 TGTACTGATCCAGCAATCTGCGCAAGGAAATGATAGAGGTTCTATTAAGGACCA 3081  
Qy 3073 GTGAAGATGCTCTTCGCAAGGCTTACAGGAAATTTCTATCCACAGCCCCCTGCTCAATTC 3132

Db 3082 GTGAAGATGCTTCTCAGGCGCTTACAGGAAATTCCTACCCACAGCCTCTTCTCAATTC 3141  
QY 3133 AGCATGTGTCTAGAGGACGACGATACACTGTCTTTTATTCATTCAGTGTCCAGTTGCCAA 3192  
Db 3142 AGCATGTGTCTGAGGACGACGATATACTAGTGTCTTTTATTCATTCAGTGTCCAGTTGCCAG 3201  
QY 3193 TGAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACC 3252  
Db 3202 TGAATGGCCACAGCAATGCAATTAATCAGAGTATCTGTGAGAAGTGTGAGAACCTGACC 3261  
QY 3253 ACAGGCAAGCACTGCAGAGCCTGCATATCTGGCTTCTACGGTATCCCAACCAATGGAGGG 3312  
Db 3262 ACAGGCAAGCACTGCAGAGCCTGCATATCTGGCTTCTACGGTATCCCAACCAATGGAGGG 3321  
QY 3313 AAATGTCAAGCATGCAAGTGCATGCGCTCTCTGTGCAACACCAACAGCGGCAAG 3372  
Db 3322 AAATGTCAAGCATGCAAGTGCATGCGCTCTCTGTGCAACACCAACAGCGGCAAG 3381  
QY 3373 TGTCTTGTGACCAACCAAGGCGTCAAGGGGACGAGTGCAGCTATGTGAGTGAAGAAAT 3432  
Db 3382 TGTCTTGTGACCAACCAAGGCGTCAAGGGGACGAGTGCAGCTATGTGAGTGAAGAAAT 3441  
QY 3433 CGATACCAAGGAAACCTCTCAGAGGACATGTTATTTATTTACTCTTCTTATTGACTATCAG 3492  
Db 3442 CGATACCAAGGAAACCTCTCAGAGGACATGTTATTTATTTACTCTTCTTATTGACTATCAG 3501  
QY 3493 TTACCTTTTGTCTATCCAGGAAGATGATCCCTATTACAGAGTATCAATTTTGTGCT 3552  
Db 3502 TTACCTTTTGTCTATCCAGGAAGATGATCCCTATTACAGAGTATCAATTTTGTGCT 3561  
QY 3553 ACTCTGACGACAAACCAAGGATTTGACATGTTTCAATGATGCTTCAAGAAATTCAC 3612  
Db 3562 ACCTCTGATGAACAAACCAAGGATTTGACATGTTTCAATGATGCTTCAAGAAATTCAC 3621  
QY 3613 CTCACATCACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCT 3672  
Db 3622 CTCACATCACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCT 3681  
QY 3673 GTTGTCTTCAAAACCAACATTAAGGAGTACAAAGATGTTTCTTAATGAGAAGTTTGTAT 3732  
Db 3682 GTTGTCTTCAAAACCAACATTAAGGAGTACAAAGATGTTTCTTAATGAGAAGTTTGTAT 3741  
QY 3733 TTTCGCAACCCCAATATCATCTTCTGTTTATGTCAGTAAATTCACCTGCCCATC 3792  
Db 3742 TTTCGCAACCCCAATATCATCTTCTGTTTATGTCAGTAAATTCACCTGCCCATC 3801  
QY 3793 AAATTCAGATGCTTCTCAGCACAGCAATTTATGACCTGGTACAGTTCTTCGTG 3852  
Db 3802 AAATTCAGATGCTTCTCAGCACAGCAATTTATGACCTGGTACAGTTCTTCGTG 3861  
QY 3853 ACTTCTTTCAGTGTCTTCTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3912  
Db 3862 ACTTCTTTCAGTGTCTTCTCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3921  
QY 3913 CAAAGTTTGTGGGCTCCAGACGCTAGAGAGCAATCTTCTCAGAGATGCCAACAGATGCC 3972  
Db 3922 CAAAGTTTGTGGGCTCCAGACGCTAGAGAGCAATCTTCTCAGAGATGCCAACAGATGCC 3981  
QY 3973 AGCGTCTCTTGTGCTCTGTAATGTCGCTTGTGAACAGATGAGGAGCTTCTGATCTT 4032  
Db 3982 AGTGTCTCTTGTGCTCTGTAATGTCGCTTGTGAACAGATGAGGAGCTTCTGATCTT 4041  
QY 4033 ATTGGGGGAGTATAAAGACTGTTTCCCAACCCCAATTCATCTGAGGCGCTGTTTGGCAAC 4092  
Db 4042 ATTGGGGGAGTATAAAGACTGTTTCCCAACCCCAATTCATCTGAGGCGCTGTTTGGCAAC 4101  
QY 4093 AAAGCGGTGCTCTCTGTTTGTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCTCT 4152  
Db 4102 AAAGCTGCTGCTCTCTGTTTGTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCTCT 4161  
QY 4153 CCTGGGAGTCAAGTCTTGTGCTGGCAGCGCTGCTGAGCATCTTCTCAGCAGATGCCG 4212

Db 4162 CCTGGGAGTCCGGCTCGCAGTGGCAGTGCCTCTGCTGACATCTTCTCAGCAGATGCCA 4221  
QY 4213 ATAGTGTACAGGAGAGAGTCCAGGAGCCGTGAGAAACCGGAGCAGCAGCCCTGTCACAG 4272  
Db 4222 GTCTGTACAGGAGAGAGTCCAGGAGCCGTGAGAAACCGGAGCAGCAGCCCTGTCGCGAG 4281  
QY 4273 CCTGGGACCTGTCATCTGA 4290  
Db 4282 CCTGGGACCTGTCATCTGA 4299

AB062913 4284 bp mRNA linear ROD 16-JAN-2002  
Mesocricetus auratus Atrn mRNA for attractin, complete cds.  
AB062913 GI:16930100  
Mesocricetus auratus (golden hamster)  
Mesocricetus auratus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
Mesocricetus.

1  
Kuramoto, T., Nomoto, T., Fujiwara, A., Mizutani, M., Sugimura, T. and  
Ushijima, T.  
Insertional mutation of the Attractin gene in the black tremor  
hamster  
Mamm. Genome 13 (1), 36-40 (2002)  
21635551  
11773967  
2 (bases 1 to 4284)  
Kuramoto, T. and Ushijima, T.  
Direct Submission  
Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center  
Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,  
Chuo-ku, Tokyo 104-0045, Japan (E-mail: tkuramoto@gan2.ncc.go.jp,  
URL: http://www.ncc.go.jp/research/rat-genome/, Tel: 81-3-3542-2511,  
Fax: 81-3-5565-1753)

FEATURES  
Location/Qualifiers  
1..4284  
/organism="Mesocricetus auratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10036"  
1..4284  
/gene="Atrn"  
1..4284  
/gene="Atrn"  
/codon\_start=1  
/product="attractin"  
/protein\_id="BAB72012.1"  
/db\_xref="GI:16930101"

AGLCLPWSRALPPPLLLLLLSLLPLPREAATAAASVAGSAAAEKACDRP  
CWNCRNPCTQCPCAGWCOCHGFRFLTGTSGGFVTDGPGNVKTKTCTWLI  
EQGNRMLEFNHFPATCSWDHLVYDGSIVAPLIAESGLIVPERDNETAPEVT  
VTSGYALLHFPDADANLTGNITYNFDMCNCSGRGCKSSNITVECESENWK  
GESCDIPCTDNCGPFHRGICNASDVRCSCSPHWQPGCSIPIVAGSPFWIREYSD  
FKLPASHKAVVNGNIMVVGGMFNSHDSWMLAYDLASREWLPNHSVNNVVRYG  
HSLALHDKIYMGKIDSTGNVTELFVHNEVLLTPKAEQYAVVGHSAHV  
TLATGRVWLVIFGHCPILYISIVQYDEKLTWSLIQTQCALVQGYGSHSVYDHR  
TKALVHGKYKAFANKYRDLVRYDVTOMATILKDSRFFRYLHTAVTSCMTLV  
FGGNTDNTSMHGAKCFSPDMAYDIACDEMSVLPBPDLDHYNRFGHSAVLHNSM  
VYFGFNSLLSDVLVFTSEQDAHSEAAACVAGPGIRCLWDTVSRCSTSWELATEB  
QAEKLSKSPKRLDHRCDQHTDCYSTANTNDCHWNCDFVNVHSCSTWELATEB  
KIDYCDPNMYCNKLTSCSCALDQNCQWEPNQCIALPENICIGWELVNSCL  
LRTAKENYDNKLSCRNHAFLATLSOKKFEVLKHLIMQSSQMSKLTTPWVG  
KRTKNYSYWCWEDMSPTNLSLQWPPSPDAGFCGLISLPSITGLKAACTINPLNG  
VCEPANSAKOCRTPCALRTACSECTSSSECCWCSNKCQVDSNAYVASFPPGCM  
BWTNPPCPNCSGYCTCSHCLBPQCGWCTDPSNTGKGCLEGSYKGLRWPQAS  
TGNYVPLPLNNSCLSDSRNWSFTHCPACQCNHCKINQISCEKCNLTITGKHE  
TCISGFTGPTNGQPCCKNCHASLNTGKFCCTTCTGAGGDEBQOLCEVNRNYG  
NPLKGCYTYLLIDYQFTTSLSQBDDRYTALNFVATPDEQNRLDMFINASKNFNLN













Best Local Similarity 89.0%; Pred. No. 0; Matches 3818; Conservative 0; Mismatches 470; Indels 0; Gaps 0;

QY 3 GGTGGCGGCGAGCGCGGCGCACTAGGCAAGGCTGAGGAGGAGCGCGCGGCGACGCGCAGC 62  
DB 72 GCGCGGCGGCGGCGGCGGCGACTGAGGCGGCTTGAAGGGGTACAGACGCGGCGACGCGCAGC 131  
QY 63 GCTCGGCGGCGAGGCGGCGGCGGCGCACTGGGACGCTGGGACGCTGACAGGCGCTGGGAGGCG 122  
DB 132 GCGCGGCGGCTGGAAGAGAGGCGAGCAGCAGCCCTGCGCGGCGGAGAGGCGGCGCTGGAGGCG 191  
QY 123 GGGGCTGGGCGGCGGCGCTGCGGCTCCCGCGGCTGCTGTCTCCACGCGCTGCGGCGCAGCGCT 182  
DB 192 GTGGCGGCGGCGGCGCTGTGTCTCCCGGCGGCTGTGTCTCGCGGCGCTATCCCGCGCGCC 251  
QY 183 GCTGCTGCTGCTGTGTGTGCTCCGCGGCGGCTGTGTGTGCTGCTGCTGCTGCTGCGGCG 242  
DB 252 GCTG 311  
QY 243 CGAGGCGGCGGCGGCGGCGGCGGCTGCGGCTGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGG 302  
DB 312 CGAGGCGGCTGCGGCTGCGGCGGCGGCTGCTGCGGCTGCGGCGGCGGCTGCGGCGGCGGCGG 371  
QY 303 TGACCGGCGGCTGCTGCTCAAGCGGCGGCTGCTGCAACCGCTGCGGCGGCGGCGGCGGCGGCGG 362  
DB 372 TGACCGGCGGCTGCTGCTCAAGCGGCGGCGGCTGCAACCGCTGCGGCGGCGGCGGCGGCGGCGG 431  
QY 363 CGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 422  
DB 432 CAGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 491  
QY 423 TTCTGGGTTGTGACAGATGACCTGGAAATATATAAATACAAAGCGAGTGCAGTGGCT 482  
DB 492 TTCTGGATTTGATACAGATGACCTGGAAATATATAAATACAAAGCGAGTGCAGTGGCT 551  
QY 483 CATTGAGGACGACCAATAGATATAGACTTCTGTTTCAATCATTTTGTCTACAGATG 542  
DB 552 CATTGAGGACGACCAATAGATATAGACTTCTGTTTCAACCAATTTTGTCTACTGATG 611  
QY 543 TAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTG 602  
DB 612 TAGCTGGGACCAATTTATGTTTATGATGGGACTCAATTTACGCACTCTGATGTGCTG 671  
QY 603 ATTTAGTGCCTCATTTGCTGAGAGAGATGGCAATGAGACTGFCCTCTGAGTGTGTC 662  
DB 672 CTTTATGCGCTCATTTGTTCTGTAAGAGATGGCAATGAGAGCGGCGGCTGGAAGTCACTGT 731  
QY 663 CACATCAGTTATGCTTGTGCTGCTATTTTGTAGTGTGCTGCTTATATTTGACTGGATT 722  
DB 732 CACTTCAGTTATGACTACTGCTATTTTTCAGTGTGCTGCTTATATCTGACTGGATT 791  
QY 723 TAATATTACTACAGTTTGTATGTGTCAAATACTGCTCAGCGCGGAGGAGTGTAA 782  
DB 792 TAATATCACTTACAAATTTGATGCTGCGCGGAAATATTTGCTCAGCGCGGAGGAGTGTAA 851  
QY 783 GATCAGTATACGCGGAACTGTGATGTGATGTGCTGAAACTGGAAGTGAAGC 842  
DB 852 GAGCAGTAAACGACGACCTGTTGAGTGTGATGTGCTGAAACTGGAAGGAGGAGTC 911  
QY 843 ATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCTCATCGAGGCACTCTGCAATTC 902  
DB 912 CTGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCTCATCGAGGCACTCTGTAACGC 971  
QY 903 AAGTATGTGAGGATGCTCTGCTTCTCAGACTGCGGCGGCTGCTGATGTTTCAAGTTC 962  
DB 972 AAGTATGTGAGGATGCTCTGCTTCTCAGACTGCGGCGGCTGCTGATGTTTCAAGTTC 1031  
QY 963 TGTACCACTAACCACTCATTTTGGACTCGAGAGCAATTTCTAACTTAAAGCTCCCGAG 1022  
DB 1032 TGTGCCAGTAAACCACTCTTTTGGACTCGAGAGCAATTTCTGACTTAAAGCTTCCAG 1091  
QY 1023 AGCATCTCATTAAGCTGTGCTCAATGGAACATTTATGTGGGTTGTTGGAGATATATGTT 1082

DB 1092 AGCTCTCTATAAAGCTGTGTGTCATTAATGTAATAATAGTGGGTTGTTGGCGGCTATATGTT 1151  
QY 1083 CAACCACTCAGATTAATAACATGTTCTAGCGTATGACCTTGTCTTAGGAGTGGCTTCC 1142  
DB 1152 CAACCACTCAGATTAATAACATGTTCTAGCGTATGACCTTGTCTTAGGAGTGGCTTCC 1211  
QY 1143 ACTAAACCGTTCTGTGAAACAATGTGGTTGATATGATGTCATCTTTGGCATATACAA 1202  
DB 1212 ACTAAACCACTTGTGAAACAATGTGGTTGATATGATGTCATCTTTGGCATATACAA 1271  
QY 1203 GGATAAATTTTACATGATGAGGAGAAATTTGATCAACCTGGGAATGTGACCAATCAGTT 1262  
DB 1272 GGATAAATTTTACATGATGAGGAGAAATTTGATCAACCTGGGAATGTGACCAATCAGTT 1331  
QY 1263 GAGAGTTTTTCAATTCATTAATCAGTCATGGGTGTTGTTGACCCCTAAAGCAAGAGCA 1322  
DB 1332 GAGAGTTTTTCAATTCATTAATCAGTCATGGGTGTTGTTGACCCCAAGCTAAGGATCA 1391  
QY 1323 GTATGAGTGGTTGGGCACTCTGCACACATGTTTACCTGGAATGGCGGAGTGTGTCAT 1382  
DB 1392 GTATGAGTGGTTGGGCACTCTGCACACATGTTTACCTGGAATGGCGGAGTGTGTCAT 1451  
QY 1383 GCTGGTCATCTTTGGTCACCTGCGCTCTCTATGATATATAAGCAATGTGACGAAATATCA 1442  
DB 1452 GTTGGTCATCTTTGGTCACCTGCGCTCTACGGATATAAAGTGTGTTGTCAGGAATATCA 1511  
QY 1443 TTTGGATGAACAACATGAGATATATACACACCGGAGGTCCTTGTGCAAGGGGTTA 1502  
DB 1512 CTTGGAAAAGAACACATGAGTATATTAACAACTCAAGGTGCTCTAGTCAAGGGGTTA 1571  
QY 1503 CGGCCATAGCAGTGTTCAGCACTAGGACCAAGGCGCTTATAGTTCACTGTTGGCTTACAA 1562  
DB 1572 TGSCCAACAGCTGTTTATGATCAGAGCAAGGCTCTGTATGTTCACTGTTGGCTTACAA 1631  
QY 1563 GCTTTTCACTGTCGAATGATGACGCTTGCAGATGATCTTACCGATATGATGTGATAC 1622  
DB 1632 GGTCTTTTCACTGTCGAATGATGACGCTTGCAGATGATCTTACCGATATGATGTGATAC 1691  
QY 1623 CCAGATGTGACCACTTCTTAAAGCAGACCGGCTTATAGTTTCCGTTACTTGCACACAGCTGTGAT 1682  
DB 1692 TCAGATGTGACCACTTCTTAAAGCAGACCGGCTTATAGTTTCCGTTACTTGCATACAGCTGTGAT 1751  
QY 1683 AGTAGTGAACAACATGCTGCTGTTTGGGGAAACACACAAATGACATCTATGAGCA 1742  
DB 1752 AGTAGTGAACAACATGCTGCTGTTGGAGGAAACACACAAATGACATCTTCCATGAGCA 1811  
QY 1743 TGGCGCAATGCTTCTTCTGATGTTTATGCGCTTATGACATGCTGTGACCGCTGCTC 1802  
DB 1812 CGGTGCAATGCTTCTTCTGATGTTTATGCGCTTATGACATGCTGTGACCGCTGCTC 1871  
QY 1803 AGTGTCTTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTTACAGAGTCTT 1862  
DB 1872 AGTGTCTTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTTACAGAGTCTT 1931  
QY 1863 ACACACAGCACATGATGTTTGGTGGTTTCAATAGTCTTCTTCCAGGCACTCT 1922  
DB 1932 GCACACAGCACATGATGTTTGGTGGTTTCAACAGCTTCTTCCAGGCACTCT 1991  
QY 1923 GGTATTTTCACTCGGAAACAGTGTGTCGCGCATCGGAGTGAAGCGCTTGTGTAGCAGCAGG 1982  
DB 1992 AGTCTTCACTCGGAGCAGTGTGATGTCACGCGAGCAGAGCTGCTTGTGTGGCAGCAGG 2051  
QY 1983 ACTGCTATTTGGTGTGTGGAAACAGGCTGTCTCAGTGTATCTCGTGGGCGCTGCGC 2042  
DB 2052 ACTGCTATTTGGTGTGTGGAAACAGGCTGTCTCAGTGTATCTCGTGGGCGCTGCGC 2111  
QY 2043 AACTCATGTAACAGAGAAAGTAAATCAGAAATGTTTTTCCAAAGAACTCTTGACCA 2102  
DB 2112 AACTGAGAAACAGAGAAAGTAAATCAGAAATGTTTTTCTTAAAGAACCTTTGACCA 2171  
QY 2103 TGACAGATGTGACCGACGACACAGATTTGTTACAGCTGTGACCGCAACCAATGACTGCA 2162  
DB 2172 TGACAGATGTGACCGACGACACAGATTTGTTACAGCTGTGACCGCAACCAATGACTGCA 2231



QY	2163	CTGGTCAATGACCAATTGTGCCCAAGGATACCCCATGTACTACTGTAAACAAGAACCCAG	2222
Db	2232	CTGGTCAATGATCACTGTGTCCCTGTGAACACACAGCTGCACAGAAGGCAGATCTCCAT	2291
QY	2223	TTTATGGTATGAAATTGCCCAAGGATACCCCATGTACTACTGTAAACAAGAACCCAG	2282
Db	2292	TGCCAATGATGACAAATTGCCCAAGGATACCCCATGTACTACTGTAAACAAGAACCCAG	2351
QY	2283	CTCAGAGACTGTGCCCTGGACACAGAACTGCCAGTGGAGCCCCGGAATCAGAGATGCAT	2342
Db	2352	CTCAGAGACTGTGCCCTAGATCAGAAATTCGCCAGTGGAACTCGAAATCAAGAGTGCAT	2411
QY	2343	TGCCCTGCCGAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAAATCATCATGTTGAA	2402
Db	2412	CGCCCTGCCGAAAACATCTGTGGCATTTGGCTGGCATTTGGTGGAAATCTCGTGTCTGAA	2471
QY	2403	AATTACTACTGCCAAGGAGAAATTATGACAAATGCTAAATTTGTTCTGAGGAACACAAATGC	2462
Db	2472	AATCACTACTGCTAAGGAAAATTATGACAATGCTAAATTTGCTCTGTAGAAACCAATGC	2531
QY	2463	CTTTTGGCTCTCTTACACCCAGAGAGGTAGAAATTTGTCCTTTAAGCAGCTGCGAAT	2522
Db	2532	CTTTTGGCTTCGTTACTGCTCCAGAGAAGGTGGAGTTTGTCTTTAAGCAGCTTCGATT	2591
QY	2523	AATGCAGTCACTCAGAGCATGTCCAAGCTCACCTTTAACCCCATGGCTGGCCCTCGGAA	2582
Db	2592	AATGCAATCATCTCAGAGCAGCTCCAAGCTCACTCTGACCTCCATGGTGTTCGGA	2651
QY	2583	GATCAATGTGTCCTACTGTGGTCTGGGAAGATATGTCCCAATTTACAATAGTTTACTACA	2642
Db	2652	GATCAATGTATCTTACTGTGTCGTGGGAGGATATGTCTCCATTCACAATAGTTTGTGCA	2711
QY	2643	GTGGATGCCCTCTGAGCCCACTGATGTGGATTTCTGTGGAATTTTATCAGAACCCAGTAC	2702
Db	2712	GTGGATGCCCATCTGAGCCCACTGATGTGGATCTGTGGGATCTTGTGCAGACCTAGTAC	2771
QY	2703	TGCGGACTGAAAGCTGCAACTCTGCATCAACCCACTCAATGTGTAGTGTCTGTGAAAGGCC	2762
Db	2772	TCGCGGACTGAAAGCTGCAACTCTGCATCAACCCACTCAATGTGAAGTGTTTGTGAAAGGCC	2831
QY	2763	TGCAAAACCACTGCTAAGCAGTGCCCGACACCATGTGCCCTTGAGGACAGCATGTGGAGA	2822
Db	2832	TGCNAACCACTGTCAGCAGTGCCTGGACACACCTGTGCCCTCGGACAGCATGTGGTGA	2891
QY	2823	TTGCAACAGCGGCACTCTGAGTGCATGTGGTGCAGCAACATGAGCAGTGTGTGACCTC	2882
Db	2892	GTGCACAGCAGCAGCTCCGAATGCATGTGGTGCAGCAACATGAGCAGTGCCTAGACTC	2951
QY	2883	CAATGCTATGTGGCCCTCCCTTCCCTTTTGGCCAGTGTATGGAAATGGTATACGATGAGCAC	2942
Db	2952	CAATGCTATGCTAGCTCCCTTCCCTTTTCCGCGCAGTGTATGGAGTGGTATACAAATGAGCAG	3011
QY	2943	CTGCCCCCTGAAAATTTGTCAGGCTACTGTACTGTAGTCACTTGTCTTGAGAGCAACAGG	3002
Db	3012	CTGCCCACTGAAAATTTGCTCTGGCTACTGTACTGTGAGCCTTGTGAGCAGGCCAGG	3071
QY	3003	CTGTGGCTGTGTACTGATCCAGCAATCTTGGCAAAAGGGAATGCAATAGAGGTTTCCCTA	3062
Db	3072	CTGTGTTGTGTACTGATCTTAGCAATATCTGGGAAGGCAATATGTATTTAGGGAGCTTA	3131
QY	3063	TAAAGGACCAAGTGAAGATGCTTCGCAAGGCCCTTACAGGAAATTTCTTATCCACAGCCCT	3122
Db	3132	TAAAGGACCCGTGAAGATGCTTCACATGCTCTTACAGGAAATGTGTACCCACAGCCCT	3191
QY	3123	GCTCAATTCACAGCATGTGCTTAGAGGACAGCAGATACAATGTGTCCTTTTCACTCACTGCC	3182
Db	3192	TCTGAACTCCACAGCATGTGCTTAGAGGACAGCAGATACAACTGTGCTCTTTTATCTCACTGCC	3251
QY	3183	AGCTTGCCAAATGCACCGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAAGTGTGA	3242
Db	3252	AGCTTGCAAGTCAATGACACAGCAAGTGCATCAACAGCAGCATCTGTGAGAAAGTGTGA	3311

[illegible]

[illegible]



1106 CAACATTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAGGGAATGGCTTCC 1165  
1143 ACTAAACCGTCTCTGGAACAATGTCGTTGTAGATGATGTCATCTTTGGCATATACAA 1202  
1166 ACTAAACCAATCTGTGAACAGTGTGGTTGTAGATATGTCATCTTTTGGCAATTACATA 1225  
1203 GGATAAAATTTACATGTATGAGGAAATTTGATCCAACTGGGAATGTGACCAATGATTT 1262  
1226 GGATAAAATCTACATGTATGAGGAAATTTGATTTCAACAGGGAACGTGACCAATGAGCT 1285  
1263 GAGAGTTTTACATTTACATTAATGAGTCATGGGTGTGTTGACCCCTAAGGCAGGAGCA 1322  
1286 GAGAGTATTTACATTTACATTAATGAGTCATGGGTGTGTTGACCCCTAAGGCAGGAGCA 1345  
1323 GTATGAGTGTGGGCACTCTGCAACATTTTACACTGGAAGATGGCCGAGTGGTCAT 1382  
1346 GTATGAGTGTGGGCACTCTGCAACATTTTACACTGGAAGATGGCCGAGTGGTCAT 1405  
1383 GCTGGTCATCTTTGGTCACTGGCTCTCTATGGATATATAGCAATGTGCGAGGAATGA 1442  
1406 GTTGGTCATCTTTGGTCACTGGCTCTCTATGGATATATAGCAATGTGCGAGGAATGA 1465  
1443 TTTGGATGAACAACATGAGTATATACACACCGAGGTGCCCTGTGCAAGGGGGTTA 1502  
1466 CTTGGAAAGAACACATGAGTATATACATACTCAGGGTGTCTTTGTGCAAGGGGGTTA 1525  
1503 CGGCCATAGCAGTGTATTACGACATAGGACCGAGGCCCTATACGTTTCATGTTGGTCAAA 1562  
1526 TGGCCACAGTAGTGTATTATGATGACAGGACCAAGGCTCTGTACGTTTCATGTTGGTCAAA 1585  
1563 GCTTTTCAGTGCATTAAGTACCGGCTTCAGATGATCTACCGCATATGATGGGATAC 1622  
1586 GGCCTTTACGCGCAACAAATACCGGCTTCAGATGATCTACCGCATATGATGGGATAC 1645  
1623 CCAGATGTGGACATCTTTAAGSACAGCGGATTTTTCCGTTACTTTGCACACAGCTGTGAT 1682  
1646 TCAGATGTGGACATCTTTAAGSACAGCGGATTTTTCCGTTACTTTGCATACAGCTGTGAT 1705  
1683 AGTGAATGAACATGCTGGTGTGTTGGGGGAAACACACATATGACATCTATGAGCA 1742  
1706 AGTGAATGAACATGCTGGTGTGTTGGGGGAAACACACATATGACATCTATGAGCA 1765  
1743 TGGCGCAAAATGCTTCTCTCAGATTTCATGGCTATGACATGGCTGTGACCGCTGGTC 1802  
1766 CGGTGCAAAATGCTTCTCTCAGATTTCATGGCTATGACATGGCTGTGACCGATGGTC 1825  
1803 AGTGTCTCCGACCTGTATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1862  
1826 AGTGTCTCCGACCTGTATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1885  
1863 ACACACAGCAGCATGTATGTTGGGTGTTTCAATAGTCTCTCCCTCAGCAGCATCT 1922  
1886 GTACACAGCAGCATGTATGTTGGGTGTTTCAATAGTCTCTCCCTCAGCAGCATCT 1945  
1923 GGTATTTCACTCGGAACAGTGTGATCGGATCGGAGTGAAGCGCTGTTTATAGCAGCAGG 1982  
1946 GGTCTTTCACTCGGACAGTGGATGCAACCGCAGTGAAGCTGTTTGTGGCAGCAGG 2005  
1983 ACTGTGTATCGTGTGTGGACACAGCGTCTCAGTGTATCTCGTGGCGCTGGC 2042  
2006 ACTGTGTATCGTGTGTGGACACAGCGTCTCAGTGTATCTCGTGGCGCTGGC 2065  
2043 AACTGTATGAACAGAGAAAGTTAAATCAGATGTTTTCAAAAGAACTCTTGACCA 2102  
2066 AACTGAAGAACAGCAGAGAAAGTTAAATCAGATGCTTTTCTAAAGAACCTTGACCA 2125  
2103 TGAAGATGTGACGACACACAGATTTTACAGTGTGACGCAACCAATGATGACCA 2162  
2126 TGAAGATGTGACGACACACAGATTTTACAGTGTGACGCAACCAATGATGACCA 2185  
2163 CTGGTCAATGACATTTGTCTCCAGGAAACACAGCTCTCAGAGGCGCATCTCCAT 2222  
2186 CTGGTCAATGACATTTGTCTCCAGGAAACACAGCTCTCAGAGGCGCATCTCCAT 2245

2223 TTTTAGTATGAGAAATGCCCCAAGGATACCCCATGTACTACTGTAAACAAGAGACCAG 2282  
2246 TGCCAAATATGAGAGTTGCCCCAAGGATACCCCATGTACTACTGTAAACAAGAGACCAG 2305  
2283 CTGACAGAGCTGTGCTCTGACACAGAACTGCGAGTGGAGCCCGGAACTCAGAGTGCAT 2342  
2306 CTGACAGAGCTGTGCTCTGACACAGAACTGCGAGTGGAGCCCGGAACTCAGAGTGCAT 2365  
2343 TGCCCTCCCGGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAAACTCATGTTTGAA 2402  
2366 CGCCCTCCCGGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAAACTCATGTTTGAA 2425  
2403 AATTACTACTGCCAAGAGAAATATGACAAATGCTAAATTTCTGTAGGAACCAATGTC 2462  
2426 AATCACTACTGCTAAGAGAAATATGACAAATGCTAAATTTCTGTAGGAACCAATGTC 2485  
2463 CTTTTCGCTCTCTTCAACACCCAGAGAGGATGATTTGCTTAAAGCAGCTGCGAAT 2522  
2486 CTTTTCGCTCTCTTCAACACCCAGAGAGGATGATTTGCTTAAAGCAGCTGCGAAT 2545  
2523 AATGACATCTCTCAGAGCATGCTCAAGCTCACCTTAAACCCCATGGGTGCGCTTCGGAA 2582  
2546 AATGCAATCATCTCAAGATATGCTCAAGCTCACCTTAAACCCCATGGGTGCGCTTCGGAA 2605  
2583 GATCAATGCTCTCTTCAACACCCAGAGAGGATGATTTGCTTAAAGCAGCTGCGAAT 2642  
2606 GATCAATGCTCTCTTCAACACCCAGAGAGGATGATTTGCTTAAAGCAGCTGCGAAT 2665  
2643 GTGATGCTCTCTTCAACACCCAGAGAGGATGATTTGCTTAAAGCAGCTGCGAAT 2702  
2666 GTGATGCTCTCTTCAACACCCAGAGAGGATGATTTGCTTAAAGCAGCTGCGAAT 2725  
2703 TCGGGGACTGAGGCTGCAACCTGATCAACCCCATCAATGCTGTAGTGTCTGTGAAGGCC 2762  
2726 TCGGGGACTGAGGCTGCAACCTGATCAACCCCATCAATGCTGTAGTGTCTGTGAAGGCC 2785  
2763 TGCAAAACACAGTGTGAAGCAGTGGGACACCATGCTTGAAGGACAGCATGTGAGGA 2822  
2786 TGCAAAACACAGTGTGAAGCAGTGGGACACCATGCTTGAAGGACAGCATGTGAGGA 2845  
2823 TTGCACACGCGGAGCTCTGATGTCATGCTGGTGGAGCAACATGAAGCAGTGTGTGACTC 2882  
2846 GTGCACCTAGCAGCAGCTCGGAGTGCATGCTGGTGGAGCAACATGAAGCAGTGTGTGACTC 2905  
2883 CAATGCTATGTGCTCTCTTCCCTTTGGCCAGTGTATGGAATGGTATACGATGAGCAC 2942  
2906 CAATGCTATGTGCTCTCTTCCCTTTGGCCAGTGTATGGAATGGTATACGATGAGCAC 2965  
2943 CTGCCCCCTGAAATTTGTTTCAAGCTACTGCTGTAGTCAATTTGCTTGGAGCAACCCAGG 3002  
2966 CTGCCCCCTGAAATTTGTTTCAAGCTACTGCTGTAGTCAATTTGCTTGGAGCAACCCAGG 3025  
3003 CTGTGGCTGTGTACTGATCCAGCAATATCTGGCAAGGGAATGATAGAGGTTCCCTA 3062  
3026 CTGTGGTGTGTACTGATCCAGCAATATCTGGCAAGGGAATGATAGAGGTTCCCTA 3085  
3063 TAAAGGACCACTGAGAGTGTGCTTCCGAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3122  
3086 TAAAGGACCACTGAGAGTGTGCTTCCGAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3145  
3123 GCTCAATTTCCAGCATGTGTCTAGAGGACAGCATCAACTGCTGCTTTTCACTACTGTCC 3182  
3146 TCTGAATCTCCAGCATGTGTCTAGAGGACAGCATCAACTGCTGCTTTTCACTACTGTCC 3205  
3183 AGCTTTCGCAATGCAACCGGACAGTAAATGTCATCAATCAGAGCATCTGTGAGAGTGTGA 3242  
3206 AGCTTTCGCAATGCAACCGGACAGTAAATGTCATCAATCAGAGCATCTGTGAGAGTGTGA 3265  
3243 GAACTGACCAACAGGACAGCTCGAGAGCTGATATCTGGCTTCTACGGTGTATCCAC 3302  
3266 GGACCTGACCAACAGGACAGCTCGAGAGCTGATATCTGGCTTCTATGGTGTATCCGAC 3325

Qy	3303	CAATGAGGGAATGTGACCATGCAAGTGCATGCGCAGCGTCTCTGTGCAACACCAA	3362
Db	3326	TAATGAGGGAATGTGACCATGCAAGTGCATGCGCAGCGTCTCTGTGCAACACCAA	3385
Qy	3363	CACGGCAAGTGTCTTGCACCAACCAAGGCGCTCAAGGGGACGAGTGCCAGCTATGTA	3422
Db	3386	CACCGCAAGTGTCTTGCACCAACCAAGGCGTCAAGGGGACGAGTGCCAGCTATGTA	3445
Qy	3423	GGTAGAAATCGATACCAAGGAACCTCTCAGAGGAACATGTTATTAATCTTCTTAT	3482
Db	3446	GGTAGAAATCGATACCAAGGAACCTCTCAGAGGAACATGTTATTAATCTTCTTAT	3505
Qy	3483	TGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGCTTATACAGCTATCAA	3542
Db	3506	TGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGCTTATACAGCTATCAA	3565
Qy	3543	TTTGTGGCTACTCTGACGAACAAACAGGATTTGACATGTTTCATCAATGCTTCAA	3602
Db	3566	CTTTGTGGCTACTCTGATGAACAAACAGGATTTGACATGTTTCATCAATGCTTCAA	3625
Qy	3603	GAATTTCAACCTCAACATCAGCTGGGCTGCGAGTTTCTCAGCTGGAACCCAGCTGGAGA	3662
Db	3626	AAATTTCAACCTCAACATCAGCTGGGCTGCGAGTTTCTCAGCTGGAACCCAGCTGGAGA	3685
Qy	3663	AGAGATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGA	3722
Db	3686	AGAGATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGA	3745
Qy	3723	GAAGTTGATTTTCGCAACCAACCAATCACTTCTTGTGTTATGTCAGTAATTTTAC	3782
Db	3746	GAATTTGATTTTCGCAACCAACCAATCACTTCTTGTGTTATGTCAGTAATTTTAC	3805
Qy	3783	CTGGCCCATCAAAATTCAGATTCCTTCTCAGACACAGCAATTTTATGGACCTGGTACA	3842
Db	3806	TTGGCCCATCAAAATTCAGATTCCTTCTCAGACACAGCAATTTTATGGACCTGGTACA	3865
Qy	3843	GTTCTTCGTGACTTTCTCAGTGTGTTCTCTCTCTTCTGCTGCTGCTGCTGCTGCTG	3902
Db	3866	GTTCTTCGTGACTTTCTCAGTGTGTTCTCTCTCTCTGCTGCTGCTGCTGCTGCTG	3925
Qy	3903	GAAGATCAAAACAAAGTTGTGGGCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCA	3962
Db	3926	GAAGATCAAGACAGAGCTGTGGGCTCCAGGCGGAGAGAGCAACTTCTTCGAGAGATGCA	3985
Qy	3963	ACAGATGGCCAGCGCT	4022
Db	3986	ACAGATGGCCAGCGCGCT	4045
Qy	4023	TCCTGATCTTATTTGGGGGAGTAAAGACTGTCTCCAAACCCATTTGACCTGGAGCGTG	4082
Db	4046	TCCTGATCTTATTTGGGGGAGTAAAGACTGTCTCCAAACCCATTTGACCTGGAGCGTG	4105
Qy	4083	TTTTGGCAACAAAGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4142
Db	4106	TTTTGGTAAACAAAGCGGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4165
Qy	4143	CATCCCTCTCTCTGGCAGTCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4202
Db	4166	AAATCCCTCTCTCTGGTCAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4225
Qy	4203	GCAGATGCTCGATGTTTCAAGGAGAGTCAAGGAGCGCTGAGAAACCCGGAAGCAGCAGCC	4262
Db	4226	GCAGATGCTCGATGTTTCAAGGAGAGTCAAGGAGCGCTGAGAAACCCGGAAGCAGCAGCC	4285
Qy	4263	CCCTGACAGCTGGGACCTGCAATCTGA	4290
Db	4286	GCCTGACAGCTGGGACCTGCAATTTGA	4313

RESULT 9

AF034957

LOCUS

DEFINITION

3597 bp mRNA linear PRI 30-SEP-1998  
Homo sapiens secreted T cell activation protein Attractin

(attractin) mRNA, complete cds.	
ACCESSION	AF034957
VERSION	AF034957.1
KEYWORDS	GI:3676346
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 3597) Duke-Cohan,J.S., Gu,J., McLaughlin,D.F., Xu,Y., Freeman,G.J. and Schlossman,S.F.
TITLE	Attractin (Dppr-L), a member of the CUB family of cell adhesion and guidance proteins, is secreted by activated human T lymphocytes and modulates immune cell interactions
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)
MEDLINE	98409658
PUBMED	9736737
REFERENCE	2 (bases 1 to 3597)
AUTHORS	Duke-Cohan,J.S., Gu,J., Ao,Z., McLaughlin,D.F., Freeman,G.J. and Schlossman,S.F.
TITLE	Direct Submission
JOURNAL	Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS, Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
FEATURES	Location/Qualifiers
source	1.. 3597
gene	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	1.. 3597
	/gene="attractin"
CDS	1.. 3597
	/gene="attractin"
	/note="secreted T cell activation protein; Method: conceptual translation with partial peptide sequencing; Dppr-L"
	/codon_start=1
	/product="Attractin"
	/protein_id="AAC61902.1"
	/db_xref="GI:3676347"
	/translation="MVAAAATEARLRRTTAATAALAGSGGPHGVGRCNPGTGCQ VCPAGVSEQCCHGGRFRLTGGTGVTPDGNFYKTKTWTLEIGQPNRIMRLRNFH FATECSMDHLYVDGDSIYAPLVAAPSGLIIPERDNETVPVATSGVALLHFFSDA AYNLGNFTITYSDFMCPNCSGRGCKISNSDSTVECESENWKEACDIPHTCNCG FPHRGICNSDVRGCGSPSDWQGPSPVPAQSFWRTEEYNSNLKLPASHKAVNG NIMVVGVMFNHSDYNNMVLAYDLASREWLPLNRSVNNVYVYHSLALYKDKIYWG KIDSTGNTNELRVPHIHESWLLTPKAQYAVVGHSAHIVLKNRNVVLMVIFG HCLPYGISNVOBYDLDKNTWSILHTGQALVQGGYGHSSVYDHRTRALYVHGKYKAFS ANKRYLADLRYVDVTOMWTILKDSRFRILHTAVIVSGTMLVFGGNTHTDTSMSG AKCFSDFWAYDIACDRWSVLPRDSTWMTSLTAIPAVLHNTMTVFGGTFNSLLSDI LVFTSQCDARHSEACLAAGGIRCVNWTGSSQISWALATDEEELKSKCECFKRT LHDRCDQHTDCTYATNDHCNDRHSCVPNHSSEGGQISIFRYPCKDNPMYIC NKRTSCRCALDQNCQWEPNQCIALPENICIGHWLVGNSCLIKITAKENYDNAL FPCRNHALLASLTQKVEFLKQIRMQSSQMSKLTLPWVGLRKINYSVWDEDM SPTNLLQWMPSEPSDAGFCILSEPSRGLKAATCINPLNGSVCEPANHSAKQCR TPCALATCGDCTSGSSECMWCSNMKQCVDSNAYVASPFQOCMEWYMTWCPPLNSM GYCTCSHCLBPQCGCTDPSNTGKGCIEGSGKGVKMPQSAPTGNFVQPLNNSM CLEDYSNWSFTHCPACQCNHSGKINQSIQCEKCNLTGKHCETCISGFGDPNGG KCQPCNKGHSLCNTNKTGCTTGKGVGSDQCEVENRYQGNPLRGTCYITLLID YQFTFSLGQEDDRYITAFNVPATPDEQNRDLDMFINASKFNLNITWAASPSAGTQAG EEMPVVSXTNIKEYKDSFSENEKDFRNHNFTFFVYVSNTFWPIKIQVTEQ"
ORIGIN	
Query Match	80.9%; Score 3469.4; DB 9; Length 3597;
Best Local Similarity	99.2%; Pred. No. 0;
Matches 3507; Conservative	0; Mismatches 26; Indels 2; Gaps 2;
Qy	270 GTCGGGCTCAGCGCAGCGGAGCCAGGAATGTACCGGCGCTGTGTCAACGGCGGTG 329
Db	48 GGCGGCGACGGCAGCGCTCGCGGCGAGGAGCGCGGCGCGACTGTGTCAACGGCGGTG 107
Qy	330 CTCAACCTTGGCACCGGCGAGTGTCTGCCCCCGCGCTGGTGGGGGAGCAATGCCA 389

Db 108 CTGCAACCCCTGGCACCGGCCAGTCGCTGCTGCCCGCGGCTGGGTGGCGGAGCAATGCCA 167  
Qy 390 GCATCGGGGGCGCGCTTACAGCTAACTGGATCTTCTGGGTCTTGTGACAGATGACACCTGG 449  
Db 168 GCATCGGGGGCGCGCTTACAGCTAACTGGATCTTCTGGGTCTTGTGACAGATGACACCTGG 227  
Qy 450 AAAATTATAATACAAAAGAGTGCACGTGGCTCAATTGAAGGACAGCAATAGAAATAAT 509  
Db 228 AAAATTATAATACAAAAGAGTGCACGTGGCTCAATTGAAGGACAGCAATAGAAATAAT 287  
Qy 510 GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGACCAATTTATATGTTTATGA 569  
Db 288 GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGACCAATTTATATGTTTATGA 347  
Qy 570 TGGGACATCAATTTATGACCGCTAGTTGCTGCATTTAGTGGCCCTCAATTGTTCTGAGAG 629  
Db 348 TGGGACATCAATTTATGACCGCTAGTTGCTGCATTTAGTGGCCCTCAATTGTTCTGAGAG 407  
Qy 630 AGATGGCAATGAGACTGCTGAGTTGTTGCCACATCAGCTTATGCTTGTGCTGCAATT 689  
Db 408 AGATGGCAATGAGACTGCTGAGTTGTTGCCACATCAGCTTATGCTTGTGCTGCAATT 467  
Qy 690 TTTTATGATGCTGCTTATAATTTGACTGGATTTAATAATTAATTAATTAATTAATTAAT 749  
Db 468 TTTTATGATGCTGCTTATAATTTGACTGGATTTAATAATTAATTAATTAATTAATTAAT 527  
Qy 750 TCCAAATACTGCTCAGCGCGAGAGAGTGAAGTCAAGTAAAGTAAAGTAAAGTAAAGTAA 809  
Db 528 TCCAAATACTGCTCAGCGCGAGAGAGTGAAGTCAAGTAAAGTAAAGTAAAGTAAAGTAA 587  
Qy 810 ATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTCTACAGACAA 869  
Db 588 ATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTCTACAGACAA 647  
Qy 870 CTGTGGTTTTCTCATCGAGGCAATCTGCAATTCAGATTCAGATGATGATGATGATGATG 929  
Db 648 CTGTGGTTTTCTCATCGAGGCAATCTGCAATTCAGATTCAGATGATGATGATGATGATG 707  
Qy 930 CTCAGACTGGAGGTCCTGGATGTTTCACTTCCCTGACAGTAACTCACTCACTTGTGAC 989  
Db 708 CTCAGACTGGAGGTCCTGGATGTTTCACTTCCCTGACAGTAACTCACTCACTTGTGAC 767  
Qy 990 TCGAGAGGAATATTTCACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGG 1049  
Db 768 TCGAGAGGAATATTTCACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGG 827  
Qy 1050 AAACATTATGCGGTTGTTGGAGGATATATGTTCAACCACTCAGATTAATACATGTTCT 1109  
Db 828 AAACATTATGCGGTTGTTGGAGGATATATGTTCAACCACTCAGATTAATACATGTTCT 887  
Qy 1110 AGGATATGACCTTGTCTTAGGAGTGGCTTCCACTAAACCTTCTGTGAACAATGTGTT 1169  
Db 888 AGGATATGACCTTGTCTTAGGAGTGGCTTCCACTAAACCTTCTGTGAACAATGTGTT 947  
Qy 1170 TGTTAGATATGTCATCTTTGGCATATATCAAGGATAAATTTACATGATTAAGAGGAAA 1229  
Db 948 TGTTAGATATGTCATCTTTGGCATATATCAAGGATAAATTTACATGATTAAGAGGAAA 1007  
Qy 1230 AATTGATCCAACTGGGAATGTACCAATGATGATGAGATTTTTCATTCATTAATGAGTC 1289  
Db 1008 AATTGATCCAACTGGGAATGTACCAATGATGATGAGATTTTTCATTCATTAATGAGTC 1067  
Qy 1290 ATGGGTGTTGTGACCCCTTAAGCAAGGAGAGTATGAGTGGTTGGGCACTCTGACACA 1349  
Db 1068 ATGGGTGTTGTGACCCCTTAAGCAAGGAGAGTATGAGTGGTTGGGCACTCTGACACA 1127  
Qy 1350 CATTTGTTACTGAAGATGGCCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1409  
Db 1128 CATTTGTTACTGAAGATGGCCGAGTGGTTCATGCTGGTTCATCTTTGGTCACTGCCCCT 1187  
Qy 1410 CTATGATATATAGCAATGTGAGGAATATGATTTGGATAAGAAACACATGAGTATATT 1469  
Db 1188 CTATGATATATAGCAATGTGAGGAATATGATTTGGATAAGAAACACATGAGTATATT 1247

Qy 1470 ACACACCCAGGTCGCCCTTGTGCAAGGGGTTTACGCCCATGACAGTGTTTACGACCATAG 1529  
Db 1248 ACACACCCAGGTCGCCCTTGTGCAAGGGGTTTACGCCCATGACAGTGTTTACGACCATAG 1307  
Qy 1530 GACACGGGCGCTTATAGCTTTCATGCTGCTGCTTACAGGCTTTTCACTGCAATTAAGTACCGGT 1589  
Db 1308 GACACGGGCGCTTATAGCTTTCATGCTGCTTACAGGCTTTTCACTGCAATTAAGTACCGGT 1367  
Qy 1590 TGCAATGATCTCTACCGATATGATGTGATACCCAGATGTGACCAATTTCTTAAGGACAG 1649  
Db 1368 TGCAATGATCTCTACCGATATGATGTGATACCCAGATGTGACCAATTTCTTAAGGACAG 1427  
Qy 1650 CCGATTTTTCGGTTACTTGCACACAGCTGTGATAGTGTGAGTGAACCATCTCTGTTTGG 1709  
Db 1428 CCGATTTTTCGGTTACTTGCACACAGCTGTGATAGTGTGAGTGAACCATCTCTGTTTGG 1487  
Qy 1710 GGAACACACACACATGACATCTATAGAGCAATGCGCGCAATGCTTCTCTTCAGATTT 1769  
Db 1488 GGAACACACACACATGACATCTATAGAGCAATGCGCGCAATGCTTCTCTTCAGATTT 1547  
Qy 1770 CATGGCTATGACATTCCTGTGACCGCTGCTGAGTGTGCTTCCAGACCTGATCTCCACCA 1829  
Db 1548 CATGGCTATGACATTCCTGTGACCGCTGCTGAGTGTGCTTCCAGACCTGAT-TCCACCA 1606  
Qy 1830 TGATGTCAAACAGATTTGGCCATT-CAGCAGTCTTTACACAAACACACATGATGTGTCG 1888  
Db 1607 TGATGTCAAACAGATTTGGCCATTCCAGCAGTCTTTACACAAACACACATGATGTGTCG 1666  
Qy 1889 GTGGTTTCAATAGTCTCTCTCAGCGACATCTGCTGATTTCACTCGAACAGTGTGATG 1948  
Db 1667 GTGGTTTCAATAGTCTCTCTCAGCGACATCTGCTGATTTCACTCGAACAGTGTGATG 1726  
Qy 1949 CGCATCGAGTGAAGCGCTTGTTTAGCAGCAGACCTGCTGATTTGCTGTGTGGAACA 2008  
Db 1727 CGCATCGAGTGAAGCGCTTGTTTAGCAGCAGACCTGCTGATTTGCTGTGTGGAACA 1786  
Qy 2009 CAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAGAGAAAGTTAA 2068  
Db 1787 CAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAGAGAAAGTTAA 1846  
Qy 2069 AATCAGATGCTTTTCCAAAGAACTCTTCACCATGACAGATGTGACACACACAGATT 2128  
Db 1847 AATCAGATGCTTTTCCAAAGAACTCTTCGACATGACAGATGTGACACACACAGATT 1906  
Qy 2129 GTTACAGCTGTACAGCCAAACCAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2188  
Db 1907 GTTACAGCTGTACAGCCAAACCAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1966  
Qy 2189 GGAACACAGCTGTGAGAGGCGAGTCTCCATTTTGTAGTATGAGAAATGCCCCAAGG 2248  
Db 1967 GGAACACAGCTGTGAGAGGCGAGTCTCCATTTTGTAGTATGAGAAATGCCCCAAGG 2026  
Qy 2249 ATAACCCCATGTACTACTGTAAACAAAGAGACACAGCTGCGAGAGCTGTGCGCTGACAGA 2308  
Db 2027 ATAACCCCATGTACTACTGTAAACAAAGAGACACAGCTGCGAGAGCTGTGCGCTGACAGA 2086  
Qy 2309 ACTGCGAGTGGAGCCCGGAAATCAGAGTGCATTTGCCCTGCGCGGAAATATCTGTGCA 2368  
Db 2087 ACTGCGAGTGGAGCCCGGAAATCAGAGTGCATTTGCCCTGCGCGGAAATATCTGTGCA 2146  
Qy 2369 TTGGCTGCTGCTGTTGGAACTCATGTTTGAATAATTTACTACTGCGCAAGAGAAATTTATG 2428  
Db 2147 TTGGCTGCTGCTGTTGGAACTCATGTTTGAATAATTTACTACTGCGCAAGAGAAATTTATG 2206  
Qy 2429 ACAATGCTAAATTTCTGTAGGAAACCAATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2488  
Db 2207 ACAATGCTAAATTTCTGTAGGAAACCAATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2266  
Qy 2489 AGAGGTGAAATTTGCTTAAAGCAGTGCAGTAAATGACATCTCAGAGCATGTGCA 2548  
Db 2267 AGAGGTGAAATTTGCTTAAAGCAGTGCAGTAAATGACATCTCAGAGCATGTGCA 2326

QY 2549 AGCTCACCTTAACCCCATGGTCCGCTTCGGAAGATCAATGTCTCTACTCTGCTGCG 2608  
Db 2327 AGCTCACCTTAACCCCATGGTCCGCTTCGGAAGATCAATGTCTCTACTCTGCTGCG 2386  
QY 2609 AAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCCCTGAGCCAGCTGATG 2668  
Db 2387 AAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCCCTGAGCCAGCTGATG 2446  
QY 2669 CTGGATCTGTGGAATTTTATCAGAACCCAGTACTCTGGGAGCTGAAGGCTGCAACTGCA 2728  
Db 2447 CTGGATCTGTGGAATTTTATCAGAACCCAGTACTCTGGGAGCTGAAGGCTGCAACTGCA 2506  
QY 2729 TCAACCCACTCAATGTAGTCTGTAAGAGCCCTGCAACACACAGCTGTAAGAGTGC 2788  
Db 2507 TCAACCCACTCAATGTAGTCTGTAAGAGCCCTGCAACACACAGCTGTAAGAGTGC 2566  
QY 2789 GGAACCACTGTGCTTGGAGACAGCATGTGGAGATTGCAACAGCGGAGCTCTGAGTGA 2848  
Db 2567 GGAACCACTGTGCTTGGAGACAGCATGTGGAGATTGCAACAGCGGAGCTCTGAGTGA 2626  
QY 2849 TGTGTGACGACATGAAGCAGTGTGGACTTCAATGCCCTATGTGGCTCTCTTCCCTT 2908  
Db 2627 TGTGTGACGACATGAAGCAGTGTGGACTTCAATGCCCTATGTGGCTCTCTTCCCTT 2686  
QY 2909 TTGGCCAGTGTATGGAATCTGTATACGATGAGCACCTGCCCTCGAAATTTGTCAGGCT 2968  
Db 2687 TTGGCCAGTGTATGGAATCTGTATACGATGAGCACCTGCCCTCGAAATTTGTCAGGCT 2746  
QY 2969 ACTGTACCTGTAGTCAATGCTTGGAGCAACAGGCTGTGGTGTGTACTGATCCAGCA 3028  
Db 2747 ACTGTACCTGTAGTCAATGCTTGGAGCAACAGGCTGTGGTGTGTACTGATCCAGCA 2806  
QY 3029 ATACTGGCAAGGAAATCATAGAGGTTTCTATAAGAGGACAGTGAAGTGCCTTCG 3088  
Db 2807 ATACTGGCAAGGAAATCATAGAGGTTTCTATAAGAGGACAGTGAAGTGCCTTCG 2866  
QY 3089 AAGCCCTTACAGGAATTTCTATCCAGCCCTCTCAATTCAGCATGTGTCTAGAGG 3148  
Db 2867 AAGCCCTTACAGGAATTTCTATCCAGCCCTCTCAATTCAGCATGTGTCTAGAGG 2926  
QY 3149 ACAGCAGATACACTGCTTCTATCTGTCAGCTGTCGATGTCGATGTCGATGTCGATG 3208  
Db 2927 ACAGCAGATACACTGCTTCTATCTGTCAGCTGTCGATGTCGATGTCGATGTCGATG 2986  
QY 3209 AATGCATCAATCAGAGCATCTGTGAGAACTGTGAGAACTTCCAGCATGTGTCTAGAGG 3268  
Db 2987 AATGCATCAATCAGAGCATCTGTGAGAACTGTGAGAACTTCCAGCATGTGTCTAGAGG 3046  
QY 3269 AGACCTGCATATCTGGCTTCTACGCTGATCCACCAATGGAGGAAATGTGAGCCATGCA 3328  
Db 3047 AGACCTGCATATCTGGCTTCTACGCTGATCCACCAATGGAGGAAATGTGAGCCATGCA 3106  
QY 3329 AGTGAATGGGACACGCTCTGTGTCACACCAACAGCGGCAAGTCTTCTGACACCA 3388  
Db 3107 AGTGAATGGGACACGCTCTGTGTCACACCAACAGCGGCAAGTCTTCTGACACCA 3166  
QY 3389 AGGGCGTCAAGGGGACAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3448  
Db 3167 AGGGCGTCAAGGGGACAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3226  
QY 3449 CTCTCAGAGGAACATGTTATTATCTCTTATTGACTATCAGTTCACCTTTAGTCTAT 3508  
Db 3227 CTCTCAGAGGAACATGTTATTATCTCTTATTGACTATCAGTTCACCTTTAGTCTAT 3286  
QY 3509 CCCAGGAAGATGATGCTTATTACAGCTATCAATTTTGTGGCTACTCTCTGACGAAACAA 3568  
Db 3287 CCCAGGAAGATGATGCTTATTACAGCTATCAATTTTGTGGCTACTCTCTGACGAAACAA 3346  
QY 3569 ACAGGGATTTGAGATGTTCAATGCTCCAGAAATTTCACTCAATCACTCTGG 3628  
Db 3347 ACAGGGATTTGAGATGTTCAATGCTCCAGAAATTTCACTCAATCACTCTGG 3406  
QY 3629 CTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTCTGTTTTCAAAACCA 3688

Db 3407 CTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTCTGTTTTCAAAACCA 3466  
QY 3689 ACATTAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTTGAATTTCCGCAACCCCAA 3748  
Db 3467 ACATTAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTTGAATTTCCGCAACCCCAA 3526  
QY 3749 ATATCACTTTCTTTTGTATTATGTCAAGTAATTTACCTGGCCCATCAAAATTCAGT 3803  
Db 3527 ATATCACTTTCTTTTGTATTATGTCAAGTAATTTACCTGGCCCATCAAAATTCAGT 3581

RESULT 10  
AR164807  
LOCUS  
DEFINITION Sequence 1 from patent US 6274339.  
ACCESSION AR164807  
VERSION AR164807.1 GI:16238010  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 8827)  
AUTHORS Moore, K. and Nagle, D. Lynn.  
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity  
JOURNAL Patent: US 6274339-A 1 14-AUG-2001;  
FEATURES Location/Qualifiers  
          1..8827  
          /organism="unknown"  
          /mol\_type="unassigned DNA"

ORIGIN

Query Match 79.2%; Score 3398; DB 6; Length 8827;  
Best Local Similarity 86.3%; Pred. No. 0;  
Matches 3869; Conservative 0; Mismatches 410; Indels 205; Gaps 2;

QY 3 GGTGGCGGACAGCGCGGCAACTGAGCAAGGCTGAGGAGGAGAGCGCGGCGGACGCGAGC 62  
Db 82 GGTGGCGGCGCGCGGCGGCGACTGAGCGCGCTGAGGGGGAGCAGAGGACGACAGCAGC 141  
QY 63 GCTCGCGGCGAGAGCGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 122  
Db 142 GCCTCGCGGCGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 201  
QY 123 GGGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 182  
Db 202 GGGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 261  
QY 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242  
Db 262 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312  
QY 243 CGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 302  
Db 313 CGAGGCGCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 372  
QY 303 TGAACCGCGCTGTGTCAACCGCGGCTGCTGCAACCGCGGCGGCGGCGGCGGCGGCGG 362  
Db 373 TGACCGCGCGTGTGTCAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 432  
QY 363 CGCGGCGTGGTGGGCGGAGCAATGCGAGCATGCGGGGGCGGCGGCGGCGGCGGCGGCGG 410  
Db 433 CACGGGCTGGTGGGCGGAGCAATGCGAGCATGCGGGGGCGGCGGCGGCGGCGGCGGCGG 492  
QY 411 ----- 410  
Db 493 ACGCCTATAATCACAGCTGTTCCGAGAGTGGAGGAGTGGAGGAGACAGTTCGAGGCAAGCTTC 552  
QY 411 ----- 410  
Db 553 GGCTACAGAATAAGTTCAAGAGTAACCTGGGCGCAACTTGGGCTTGTCTCCCAAAACCAAAA 612

QY 411 ----- 410  
Db 613 TGAGCGAAAGGAGCAAGCTAGAGTCTTTTGGGAAATTTTAGCTGACTAAATTTTCCACC 672  
QY 411 ----ACTAACTGGATCTTCTGGGTTTGTGACAGATGGAACCTCGAAATTTATAATACAAA 466  
Db 673 GAGNACTAACTGCTCTTCTGGATTTGTACAGATGGAACCTGGGAATTTATAATATAAGA 732  
QY 467 CGAAGTCAGCTGGCTCAATTGAAGGACAGCCAAATAGATAATAGAGACTTGTATCAATC 526  
Db 733 CGAAGTCAGATGGCTCAATTGAAGGACAGCCAAATAGATAATAGAGACTTGTATCAATC 792  
QY 527 ATTTTGTCTACAGAGTGTAGTGGGACCAATTATATGTTTATGATGGGACTCAATTTATG 586  
Db 793 ATTTTGTCTACAGAGTGTAGTGGGACCAATTATATGTTTATGATGGGACTCAATTTATG 852  
QY 587 CAGCGTGTGTGATTTAGTGGGCTCATTTGCTTCTGAGAGAGATGGCAATGAGACTG 646  
Db 853 CAGCTGTGATTTGCTGCTTTAGTGGGCTCATTTGCTTCTGAGAGAGATGGCAATGAGAGCG 912  
QY 647 TCCTGAGGTTGTGGCAGATCAGGTTATGCCCTTGTGCTGCATTTTATGATGGGACTCAATTTATG 706  
Db 913 CTCTGAGGTCAGTGTCACTTCAAGTTATGCACTGCTGCTGCTTTTTCAGTGAAGTGTCTT 972  
QY 707 ATAATTTGACTGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 766  
Db 973 ATAATTTGACTGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1032  
QY 767 GCCGAGGAGTGTAGATCAGTAAATAGCAGGCAACTGTTGAAATGTAATGTTCTGAAA 826  
Db 1033 GCCGAGGAGTGTAGATCAGTAAATAGCAGGCAACTGTTGAAATGTAATGTTCTGAAA 1092  
QY 827 ACTGGAAGAGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTCTCATC 886  
Db 1093 ACTGGAAGAGGAGTGTGACATTCCTCACTGTACAGACAACTGTGGTTTCTCATC 1152  
QY 887 GAGGATCTGCAATTCAGTGTGTGAGGATGCTTCTGCTTCTCAGACTGGCAGGCTC 946  
Db 1153 GAGGATCTGTAATGACAGGATACAGAGGTTGCTTCTGCTTCTCAGCTGGCAGGCTC 1212  
QY 947 CTGATGTTCTGATTCCTGACCACTCACTGATTAACATGTTTGGACTCGAGAGAAATTTCTA 1006  
Db 1213 CTGATGTTCTGATTCCTGACCACTCACTGATTAACATGTTTGGACTCGAGAGAAATTTCTG 1272  
QY 1007 ACTTAAAGCTCCCGAGGATCTCATAAAGCTGTGGTCAATGGAACATTTATGTTGGTTG 1066  
Db 1273 ATTTAAAGCTCCCGAGGCTCTCAATAAGCTGTGGTCAATGGAACATTTATGTTGGTTG 1332  
QY 1067 TTGAGGATATATGTTCAACCACTCAGATTAATAACATGTTTCTAGCGTATGACCTTGTCT 1126  
Db 1333 TTGAGGATATATGTTCAACCACTCAGATTAATAACATGTTTCTAGCGTATGACCTTGTCT 1392  
QY 1127 CTAGGAGTGGCTTCCACTAAACCTTCTGTGAACTATGTTGTTTGTAGATATGTTCTAT 1186  
Db 1393 CTAGGAGTGGCTTCCACTAAACCTTCTGTGAACTATGTTGTTTGTAGATATGTTCTAT 1452  
QY 1187 CTTTGGCATTTACAAAGGATAAAATTTACATGTATGAGAGAAAAATTTGATCCAACTGGGA 1246  
Db 1453 CTTTGGCATTTACAAAGGATAAAATTTACATGTATGAGAGAAAAATTTGATCCAACTGGGA 1512  
QY 1247 ATGTGACCAATGATTTGAGAGTTTTCATTTCAATTAATGAGTCATGTTGTTGTTGACCC 1306  
Db 1513 ACCTGACCAATGAGCTGAGAGTTTTCATTTCAATTAATGAGTCATGTTGTTGTTGTTGAC 1572  
QY 1307 CTAAGGCAAGGAGCAGTATGAGTGTGTTGGGCACTCTGCAACATTTGTTACACTGGAAGA 1366  
Db 1573 CGAAAGCTAAGGATCAGTATGAGTGTGTTGGCACTCTGCAACATTTGTTACACTGGAAT 1632  
QY 1367 ATGCGGAGTGTGATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1426  
Db 1633 CTGGCGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1692  
QY 1427 ATGTGAGGAAATATGTTTGGATAAGAACACATGAGTATATTAACACACCCAGGTTGCC 1486

Db 1693 TTGTGAGGAATATGATTTGGAAGAAACACATGAGTATATTAACATCTCAGGTTGCTC 1752  
QY 1487 TTGTGCAAGGGGTTTACGGCCATAGCAGTGTATTAAGACCATAGGACAGGAGGCTTATACG 1546  
Db 1753 TTGTGCAAGGGGTTTATGSCCAGTGTATTAAGACCATAGGACAGGAGGCTTATACG 1812  
QY 1547 TTCAATGGCTTACAAAGGCTTTTCAAGTGTGCAATTAAGTACGGGCTTCAAGTGTCTTACC 1606  
Db 1813 TTCAATGGCTTACAAAGGCTTTTCAAGTGTGCAATTAAGTACGGGCTTCAAGTGTCTTACA 1872  
QY 1607 GATATCATGTGGATACCCAGATGTGAGCATTCTTTAAAGACAGCCGATTTTTCGGTTACT 1666  
Db 1873 GATACATGTGGATACCCAGATGTGAGCATTCTTTAAAGACAGCCGATTTTTCGGTTACT 1932  
QY 1667 TGACACAGCTGTGATGAGTGGAAACATGCTGGTGTGTTGGGGGAAACACACAAATG 1726  
Db 1933 TGACACAGCTGTGATGAGTGGAAACATGCTGGTGTGTTGGGGGAAACACACAAATG 1992  
QY 1727 ACACATCTATGAGCCATGCGCCAAATGCTTCTTTCAGATTTTCATGCGCTTATGACATG 1786  
Db 1993 ACATTTCCATGAGCCAGGTTGCCAATGCTTCTTCCGACCTTATGACATG 2052  
QY 1787 CTGTGACCGCTGTGATGAGTGTGCGACACCTGATCTCCACATGATGTCAACAGATTTG 1846  
Db 2053 CTGTGACCGCTGTGATGAGTGTGCGACACCTGATCTCCACATGATGTCAACAGATTTG 2112  
QY 1847 GCCATTTACAGAGTCTTTACACACAGCACCATGATGTTGGTGGTGGTTCATATGCTTCC 1906  
Db 2113 GCCATTTACAGAGTCTTTACACACAGCACCATGATGTTGGTGGTGGTTCATATGCTTCC 2172  
QY 1907 TCCTCAGCAGATCTGTGATTTACCTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCG 1966  
Db 2173 TCCTCAGTGAACCTTGTGTTTACCTCGGAGCAGTGGCGTGCACACCCGAGTGAAGCTG 2232  
QY 1967 CTTGTTTACAGCAGGACCTGTGTTTTCGGTGTGTTGGAACACAGGCTGCTCTAGTGTG 2026  
Db 2233 CTTGTTGTCAGCAGGACCTGTGTTTTCGGTGTGTTGGAACACAGGCTGCTCTAGTGTG 2292  
QY 2027 TCTGCTGGGCGTGGCACTGTGATGAACAGAAAGAAAGTTTAAATCAGAAATGTTTTCCTA 2086  
Db 2293 CTTCTCTGGGAGTTGGCACTGGAAGAACAGCAGAAAGTTTAAATCAGAGTGTTTCCTA 2352  
QY 2087 AAAGAACTCTTGACCATGACATGTGACAGCAGCAGATTTGTATACAGCTGTACAGCA 2146  
Db 2353 AAAGAACTCTTGACCATGACATGTGACAGCAGCAGATTTGTATACAGCTGTACAGCA 2412  
QY 2147 ACACCAATGATGCCACTGGTGCATGACCAATGACCAATGTTGTTCCAGAGAACACAGCTGCTCAG 2206  
Db 2413 ATACCAATGATGCCACTGGTGCATGACCAATGACCAATGTTGTTCCAGAGAACACAGCTGCTCAG 2472  
QY 2207 AAGGCCAGATCTCCATTTTGTGATGAGATTTGCCCAAGGATAACCCATGTTACTACT 2266  
Db 2473 AAGGCCAGATCTCCATTTTGTGATGAGATTTGCCCAAGGATAACCCATGTTACTACT 2532  
QY 2267 GTAAAGAGAACACAGCTGTCAGGAGCTGTGCCCTGGACAGCAACTGCCAGTGGGAGCCCC 2326  
Db 2533 GCAATTAAGAAACACAGCTGTCAGGAGCTGTGCCCTAGACAGCAACTGCCAGTGGGAGCCCC 2592  
QY 2327 GGNATCAGAGTGTGATTTGCCCTGCCGAAATATCTGTGSCATTTGGCTTGGTTG 2386  
Db 2593 GGNATCAGAGTGTGATTTGCCCTGCCGAAATATCTGTGSCATTTGGCTTGGTTG 2652  
QY 2387 GAACTCATGTTTGAATTTACTTGTGCAAGGAGATTTATGCAATGCTAAATTTGTTCT 2446  
Db 2653 GAACTCATGTTTGAATTTACTTGTGCAAGGAGATTTATGCAATGCTAAATTTGTTCT 2712  
QY 2447 GTAGGAACCAATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2506  
Db 2713 GTAGGAACCAATGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2772  
QY 2507 TTAAGCAGCTGGCAATTAATGACAGTCTCATGAGCATGTCCAAGCTCACTTAACCCCAT 2566

Db	2773	TTAAGCAGCTTCGATTAAATCGAAATCATCTCAAGATGATGTCCAAGCTCAGTCTGCATCTCAT	2833
Qy	2567	GGGTGGCGCTTCGGAAGATCAATGTGTCTACTGTGTCTGGTGAAGATATGTGCCCATTTA	2626
Db	2833	GGGTGGTCTTCGGAAGATCAATGTGTCTACTGTGTCTGGAGGATATGTCTCCATTCA	2892
Qy	2627	CAATATGTTTACTACAGTGGATGCCGTCTGAGCCCACTGATGATGCTGGGAATTT	2686
Db	2893	CAATATGTTTCTCTCAGTGGATGCCATCTGAGCCCACTGATGCTGGCTCTGTGGGATCT	2952
Qy	2687	TATCAGAAACCCAGTACTCGGGCACTGAAGGCTGCAACTCTGCATCAACCCACTCAATGGTA	2746
Db	2953	TGTCAAGCCTTAGTACTCGGGGATTAAAGCTGCAACTCTGCATCAACCCCTCTCAATGGCA	3012
Qy	2747	GTGTCTGTGAAGCCCTGCAAAACACAGTGTAAAGCAGTCCCGGACACCAATGTCCCTTGA	2806
Db	3013	GCGTCTGTGAAGCCCTGCAAAACACAGTGTCAAGCAGTCCCGACACCAATGTCCCTGTC	3072
Qy	2807	GGACAGCATGTGGAGATTGCACCAAGCGCAGCTCTGAGTGCAGTGTGGTGCAGCAACATGA	2866
Db	3073	GGACAGCGTGTGGCAGTGTCACTAGCAGCAGCTCGAGTGCATGTGGTGCAGTAAACATGA	3132
Qy	2867	AGCAGTGTGGACTCCAAATGCCCTATGTGCCCTCTTCCCTTTTGGCCAGTGTATGGAAAT	2926
Db	3133	AGCAGTGTGTGGACTCCAAATGCCCTACGTGGCCCTCTTCCCTTTTGGCCAGTGTATGGAAAT	3192
Qy	2927	GGTATACGATGAGCAGCTGCCGCCCTGAAATTTGTTACGCTACTGTACTGTAGTCAATTT	2986
Db	3193	GGTATACGATGAGCAGCTGCCGCCCTGAAATTTGCTCTGGCTACTGTACTGTGAGCAATTT	3252
Qy	2987	GCTTTGGAGCAACAGGCTGTGGCTGTGTACTGATCCCAAGCAATACTGGCAAGGGAAT	3046
Db	3253	GCTTGGAGCAGCAGGCTGTGGTTGGTGTACTGATCTTAGCAATACTGGGAAGGAAAT	3312
Qy	3047	GCATAGAGGTTTCTATAAAGACAGTGAAGATGCTTCGCAAGCCCTTACAGGAAATTT	3106
Db	3313	GTATTAGGGCGAGCTATAAAGSACCTGTGAAGATGCCGTACAGGGCTCTGCAGGAAATG	3372
Qy	3107	TCTATCCACAGCCCTGCTCAATCCAGCATGTCTAGAGGACAGCAGATCAACTGGT	3166
Db	3373	TGTATCCACAGCCCTTCTGAATCCAGCATGTGTCTAGAGGACAGCAGATCAACTGGT	3432
Qy	3167	CTTTTCATTTACTGTCCAGCTTGCCAATGCAACGGCCACAGTAAATGCAATCAGACCA	3226
Db	3433	CTTTTCATTTACTGTCCAGCTTGCAGTGCAAACGGACACAGCAAAATGCAATCAACAGAGTA	3492
Qy	3227	TCTGTGAGAAGTGTGAGAACCTTGACCACAGGAAGCACTGCCAGACCTGCATATCTGGCT	3286
Db	3493	TCTGTGAGAAGTGTGAGGACCTTGACCACGGGCAGCACTGCCAGACCTTGCAATCTGGCT	3552
Qy	3287	TCTACGGTGATCCCAACAAATGGAGGGAATGTACGCCATGCAAGTGCATATGGGCAACGGT	3346
Db	3553	TCTATGGTGACCCGACTAATAGGAGCAAAATGTACGCCATGCAAGTGCATATGGGCAACGGT	3612
Qy	3347	CTCTGTGCAACCAACACAGGCAAGTGTCTTGACCACCAAGGGCTCAAGGGGGACG	3406
Db	3613	CACGTGTGCAACCAACACCGGCAAGTGTCTGTACCAACCAAGGTTCAAGGGGGACG	3672
Qy	3407	AGTGCCAGCTATGTGAGGTAGAAAATTCGATACCAAGGAACCCCTCTCAGAGGAAATCTGTT	3466
Db	3673	AGTGCCAGCTATGTGAGGTAGAAAATTCGATACCAAGGAACCCCTCTCAGAGGAAATCTGTT	3732
Qy	3467	ATTATACTCTTCTTATTTGGACTATCAGTTACCTTTAGTCTATCCACAGGAAGATGATCGCT	3526
Db	3733	ACTATACCCCTTCTCATTTGACTATCAGTTACCTTTAGTCTGTCCACAGGAACGACCGCT	3792
Qy	3527	ATTACACAGCTATCAATTTTGGGTACTCTCTGACGAAACAAACAGGCAATTTGGACATGT	3586
Db	3793	ACTACACAGCCATCAACTTTTGGCTACTCTCTGTATGAACAAACAGGCAATTTGACATGT	3852
Qy	3587	TCATCAATGCCCTCGAAGAAATTTCAACTCAACATCACTGGGCTGCCAGTTTCTCAGCTG	3646
Db	3853	TCATCAATGCCCTCCAAAACCTTCAACCTTCAACATCACTGGGCAACCAAGCTTCCAGCGC	3912

QY	3647	GAACCCAGGCTGGAGAGAGATGCCCTGTGTTTTCAAAAACCAACATTAAAGGAGTCAAAAG	3706
DB	3913	GAACCCAGACTGGAGAGAGGTGCCCTGTGTTTTCAAAAACCAACATCAAGGAATCAAAAG	3972
QY	3707	ATAGTTTCTCTAATGAGAAAGTTTGATTTTCGCAACCAACCAATAATACATTCTCTTTGTTT	3766
DB	3973	ATAGCTTCTCTAATGAGAAATTTGATTTTTCGCAACCATCCRAACATCACATTCTTTGTTT	4032
QY	3767	ATGTCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCCTTCTCTCAGCACAGCAAT	3826
DB	4033	ATGTGAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCCTTCTCCAGCACAGCAACT	4092
QY	3827	TTATGGAACCTGGTACAGTTCTTTCTGTGACATTTCTTCAGTTGTGTTTCCCTCTCTTTGCTTCTGG	3886
DB	4093	TCATGGAACCTGGTACAGTTCTTTCTGTGACATTTCTTCAGTTGTGTTTCCCTCTCGTGCTTCTGG	4152
QY	3887	TGCTCTGTGTGTTTGGAAAGATCAACAAGATTTGTGGCCCTCCAGACGTAGAGAGCAAC	3946
DB	4153	TGCTCTGACTGTCTGGAAGATCAACAGACAGCTGTTTGGCATCCAGGCGGAGAGCAAC	4212
QY	3947	TTCTTCGAGAGATGCAACAGATGGCGACGGTGCCTTTTGCCCTCTGTAAATGTGCGCTTGG	4006
DB	4213	TTCTTCGGGAGATGCAACAGATGGCGACGGCGCCCTTTTGCTTCTGTAAACGTGTGCTTGG	4272
QY	4007	AAACAGATCAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCTCCAAACCCA	4066
DB	4273	AAACAGATCAGGAGCCTCCTGATCTTATTGGGGGGAGTATAAAGACTGTTCCTCCAAACCCA	4332
QY	4067	TTGCACTGAGGCGGTGTTTGGCAACAAAGCGGTGCTCTCTGTGTGTTGTGAGGCTCC	4126
DB	4333	TTGCACTGAGGCGGTGTTTGGCAACAAAGCGGTGCTCTCTGTGTGTTGTGAGGCTCC	4392
QY	4127	CTCAGGCGCTGGGTGGCATCCCTCCTCTCTGGGCAGTCAAGTCTTGTCTGTGCCACAGCGCC	4186
DB	4393	CTCAGGCGCTGGGTGGCATCCCTCCTCTCTGGGCAGTCAAGTCTTGTCTGTGCCACAGCGCC	4452
QY	4187	TGTTGACATTTTCTCAGCAGATGCCGATAGTGTACAGAGAAAGTCAAGGACCGCTGAGAA	4246
DB	4453	TGTTGACATTTTCTCAGCAGATGCCGATAGTGTACAGAGAAAGTCAAGGACCGCTGAGAA	4512
QY	4247	ACCGAAGCAGACGCCCTCTGACAGCGCTGGGAOCTGCATCTGA	4290
DB	4513	ACCGAAGCAGACGCCCTCTGACAGCGCTGGGAOCTGCATCTGA	4556

RESULT 11	AB038388	4500 bp	mRNA	linear	ROD 18-JAN-2001
LOCUS	AB038388				
DEFINITION	Rattus norvegicus Attrn mRNA for attractin, complete cds.				
ACCESSION	AB038388				
VERSION	AB038388.1	GI:12275311			
KEYWORDS	attractin.				
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Rattus norvegicus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus...				
REFERENCE	1 (sites)				
AUTHORS	Kuramoto, T., Kitada, K., Inui, T., Sasaki, Y., Ito, K., Hase, T., Kawaguchi, S., Ogawa, Y., Nakao, K., Barsh, G.S., Nagao, M., Ushijima, T. and Serikawa, T.				
TITLE	Attractin/mahogany/zitter plays a critical role in myelination of the central nervous system				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)				
MEDLINE	21143347				
PUBMED	11209055				
REFERENCE	2 (bases 1 to 4500)				
AUTHORS	Kuramoto, T., Serikawa, T. and Ushijima, T.				
TITLE	Direct Submission				
JOURNAL	Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center Research Institute, Carcinogenesis Division; Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan [E-mail:tkuramoto@ncc.go.jp,				



Tel: 81-33542-2511 (ex. 4521). Fax: 81-35565-1753)

## FEATURES

```

source
1. 4500
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="WTC"
/db_xref="taxon:10116"
gene
1. 4500
/gene="Atrn"
CDS
61..388
/gene="Atrn"
/codon_start=1
/product="attractin"
/protein_id="BAB21018.1"

```

gene

/gene="Atrn"  
61.3888  
CDS

## CDS

/>\_translati="G1:17247453312"  
 WPRAGLCPRVLGRALSPPLPLPLPLPLPLPREABASAAVAASGSAABAAC  
 BCDFPVNGCNGQVCQVGTWGVGQCGOGRFLRGTSSGVTGDSGSAAYK  
 CMTLIEQPNKMLRLFAHTEKSSWDHLYIYVDGSIYAPLIAAFGLSVIPERDNGC  
 APEVYTSYALLHFTFSDAANLTFGNITVYDMSGPCNNCSGKSGSSVTECE  
 SENWKGSSCDIPIHCTDNCQPHRGIICNASDTVGCSPFHMOQSGTIPVANSQFWTR  
 FEYSDILPRASHACVINGNIMVGVGVFNHSDYSWMLAYDLASREWLISNHSVNS  
 VVRYGHSALHKKOIKYWGCKIDSTGVNTNELRVPHINHSWLLTPKADQVAVGH  
 SAHILUSSGRVNLVIFGHCPILGYIVSYQBYEDLKEHTNLSICTOQALVQGVGVHS  
 VYHTRTKLVYGYZAFPSKRLADOLIRYDVIDTQWTLTKOSRFRYLHTATVLS  
 GTMLVFGNTNLSHSGAKFSCDFMAYD.ACDRWYSRPELHLDVNRFGHSVAL  
 HNSYVYFGFNSLISLHGAFTSCSDAHRSGASCAVAGPIRLDMVTSRRTSWE  
 QIATKQKLCRSCFKRITLHDHCRDQHTQDCTSTANDNGCHDCEVPSHTG  
 LSIATKNDPCNPKMYCNKTKSCRSALDNCQWEPNDCGECALPENTCGI.GMLHY  
 GNSCLTKITAKENYDNAKLSRHNHAFSLATSSQKQVFLKRLMQSSQSKULT  
 TPMVJLKIINSYKVCWEDMSPTNLSLQWSPSESDAGFCGLSEPTGTRKAATCIN  
 PLNGSVVERPANHSACQCPALPTACCECTSSSECCWCSNNKQCDVNSAVASEP  
 FSHAGTENVYTPQLNNSGKCPDNRYNWSFTHCPACQDNGHSKINGSIKEKEDLT  
 GKCECTISGFYDPTNGKQCPCKNGHSLCINTNTKQCTCTGVGKCECQLINAK  
 NRCQNPJLKECTCYTLLIDYFTLSQBEDRYITAFVINFATPQDNRLDMFLNASK  
 NFNLNITWATSFAGTQTBEPVAVSWKTIKEYKQSFNCSFDPFNPHNTFFVYVINS  
 TPTWTKTQVRS"

ORIGIN

Query Match	72.5%	Score 3111.6	DB 10	Length 4500	
Best Local Similarity	88.7%	Prod. No. 0			
Matches 3369	Conservative	0	Mismatches 429	Indels 0	Gaps 0
QY	3	GGTGGCCCGACGGCGGCAACTGAGGCAAGGCTGAGGAGGAGACGGCGGCACGGGACG	62		
Db	72	GGGGCGCGCGGAGGCGACTGAGGCGCGTTGAGGGGTTACACGACGGCGACAGCAGC	131		
QY	63	GCTGCGGGCAGGACGCGCGGCGCCACTGGGACTGGGACGTGGACCTGGACAGGGCTGGGAGGCC	122		
Db	132	GCCCGCGGCTGGGAAGGAGGCGACACCGACCTGCGCCCGCACAGGGGGCTGGAGGCC	191		
QY	123	GGGGCTGGGGCGCGGCTGCGCTCCCGGGGTGCTGTCTCCACCGCTGCGGGCCACCGCT	182		
Db	192	GTGGCGCGCGCGGGCTGTGTCTCCGGGGTGTGTCCGGGGGCTATCCCGCGCGCC	251		
QY	183	GGTGTGTGTGTGTGTGTCTCCCGCGCGCGTGTGTCTGTCTGTCTGTCTGTCTGTGAGGC	242		
Db	252	GCTGTGCGCGTGTCTGCGCGCTCTCTCTTTTCGCTATTGTGTGCCACTGCGCCCGGGAGGC	311		
QY	243	CGAGCGCGCGGGCGGGCGGGCGGTGTGCGGCTCAGCCGCGAGCGGAGGCCAAGGAATG	302		
Db	312	CGAGCGCGCTGCGGTGGCGGGCGGCTGTGTGAGGCTCGGCCGAGCTGAGGCCAAGGAATG	371		
QY	303	TGACCGGCGCTGTGTAAACGGCGGTGTGTGTGCAACCTGTGGCACCGGCCAGTGGGTTGGCC	362		
Db	372	TGACCGGCGGTGTGTCAACGGCGGCGGTGTGCAACCTGTGGCACCGGCCAGTGGGTTGGCC	431		
QY	363	CGCGCGCTGGGTGGGCGAGCAATGCAGCACTGCGGGGGCGGCTTCAGACTAACTGGGATC	422		
Db	432	CACGGGCTGGGTGGGCGGCAATGCCAGCACTGCGGGGGCGGCTTCAGACTAACTGGATC	491		
QY	423	TTCTGGGTTTGTGACAGATGGACCTGGAAATTTAATAACAAACGAGATGCACGTGGCT	482		

Db 1572 TGCCACAGCAGTGTATTATGATCACAGGACCAAGGCTCTGTATGTTTATGTTGCTGTACAA 1631  
QY 1563 GGCTTTTCAAGTACCAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTTGATAC 1622  
Db 1632 GGCTTTTCAAGTACCAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTTGATAC 1691  
QY 1623 CAGATGTGACCAATCTTAAAGGACAGCGATTTTTCGGTTACTTGTGACACAGCTGTGAT 1682  
Db 1692 TCAGATGTGACCAATCTTAAAGGACAGCGATTTTTCGGTTACTTGTGACACAGCTGTGAT 1751  
QY 1683 AGTGAAGTGAACCAATCTGTTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCA 1742  
Db 1752 AGTGAAGTGAACCAATCTGTTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCA 1811  
QY 1743 TGGCGGCAAAATGCTTCTCTCAAGTTTCATGTCCTATGACATTTGCTGTGACCGCTGGTC 1802  
Db 1812 CGGTGCAAAATGCTTCTCTCAAGTTTCATGTCCTATGACATTTGCTGTGACCGCTGGTC 1871  
QY 1803 AGTGTCTTCCAGACCTGATCTCCACATGATGTCACAGATTTGGCCATTCAGCAGTCTT 1862  
Db 1872 AGTCTTCCAGACCTGATCTCCACATGATGTCACAGATTTGGCCATTCAGCAGTCTT 1931  
QY 1863 ACACAAACAGCAGCATGATGTTGGTGGTTTCAATAGTCTCTCTCAGCGACATCCT 1922  
Db 1932 GCACAAACAGCAGCATGATGTTGGTGGTCTTCAACAGCCTCTCTCAGTGAAGTCTT 1991  
QY 1923 GGTATTCCTCGGAACAGTGTGATCGGATCGGAGTGAAGCCGCTGTTTGAAGCAGG 1982  
Db 1992 AGTCTTCACTCGGAGCAGTGTGATGCAACCGCAGCGAAGCTGTTGTGGCAGCAGG 2051  
QY 1983 ACTGTGATTTGGTGTGTGGAACACAGGGTCTCTCAAGTGTATCTCGTGGGGCTGGC 2042  
Db 2052 ACTGGCATCCGGTGTCTGTGGGACACACAGTCTCTCGATGTACTCTGGAGTTGGC 2111  
QY 2043 AACTGATGAACAGGAAGAAAGTTAAATCAAGATGTTTTTCCAAAGAACTCTTTGACCA 2102  
Db 2112 AACTGAAGAACAGCAGAAAGTTAAATCAAGATGTTTTTCCAAAGAACTCTTTGACCA 2171  
QY 2103 TCACAGATGTGACAGCAGACACAGATTTGTACAGCTGTACAGCCACACCAATGATGCTCCA 2162  
Db 2172 TCACAGATGTGACAGCAGACACAGATTTGTACAGCTGTACAGCCACCAATGATGATGCTCCA 2231  
QY 2163 CTGGTGCAATGACCAATGTGTTCCCGAGGAACACAGCTGTCTCAGAGGCCAGATCTCCAT 2222  
Db 2232 CTGGTGCAATGACCAATGTGTTCCCGAGGAACACAGCTGTCTCAGAGGCCAGATCTCCAT 2291  
QY 2223 TTTTAGGTATGAGATTTGCCCGAGGATAACCCATGTACTACTGTAAACAGAGACAG 2282  
Db 2292 TGCCAGATGACCAATTGCCCGAGGATAACCCATGTACTACTGTAAACAGAGAACCCAG 2351  
QY 2283 CTGACAGAGCTGTGCCCTGGACAGAACTGCCAGTGGGAGCCCGGAATCAGGAGTGCAT 2342  
Db 2352 CTGACAGAGCTGTGCCCTGGATAGATCAGAACTGCCAGTGGGAACTCGAATCAGAGTGCAT 2411  
QY 2343 TGCCCTGCGGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACTCATGTTTGA 2402  
Db 2412 CGCCCTGCGGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACTCATGTTTGA 2471  
QY 2403 AATTACTACTGCCAGGAGAAATATGACAAATGCTTAAATTTGTTCTGTAGGAACCAATGC 2462  
Db 2472 AATCACTACTGTAAAGGAAATATGACAAATGCTTAAATTTGTTCTGTAGGAACCAATGC 2531  
QY 2463 CCTTTTGGCTTCTTTCAACCCAGAGAGAGTGAATTTGTCTTAAGCAGCTGCGAAT 2522  
Db 2532 CTTTTTGGCTTCTTTTACCTCCAGAGAGAGTGAATTTGTCTTAAGCAGCTTGCAT 2591  
QY 2523 AATGAGTCTCTCAGAGCATGTCACAGCTCACTTAACCCATGCTGGCTGGCTTGGAA 2582  
Db 2592 AATGCAATCATCTCAGAGCAGCTCAAGCTCACTCTGACTCCATGGGTGTGTTCTTGGAA 2651  
QY 2583 GATCAATGTGCTTACTGGTCTGGGAGATATGCTCCCAATTTTACAAATAGTTTACTACA 2642  
Db 2652 GATCAATGTATCTTACTGGTCTGGGAGATATGCTCCCAATTTTACAAATAGTTTACTACA 2711

QY 2643 GTGATGCCCTGTGAGCCAGTGTGATTTCTGTGGAAATTTTATCAGAAACCCAGTAC 2702  
Db 2712 GTGATGCCATCTGAGCCAGTGTGATGCCGCTTCTGTGGATCTTGTGAGGCTTAGTAC 2771  
QY 2703 TCGGGGACTGAAGGCTCAACCTGTGCATCAACCCACTCAATGGTGTGTCTGTGAAAGGCC 2762  
Db 2772 TCGGGGACTGAAGGCTCAACCTGTGCATCAACCCACTCAATGGTGTGTCTGTGAAAGGCC 2831  
QY 2763 TGAACACACAGTGTGAAGGCTGCGGACACCATGTGCTTGAAGACAGCATGTGAGGA 2822  
Db 2832 TGAACACACAGTGTGAAGGCTGCGGACACCATGTGCTTGAAGACAGCATGTGAGGA 2891  
QY 2823 TTGACACAGGCGGAGCTCTGAGTGCATGTGCTGCAGCAACATGAAGCAGTGTGTGACTC 2882  
Db 2892 GTGACACAGCAGCAGCTCCGAATGTCATGTGCTGCAGCAACATGAAGCAGTGTGTGACTC 2951  
QY 2883 CAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTTATGATGAGCAC 2942  
Db 2952 CAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTTATGATGAGCAC 3011  
QY 2943 CTGCCCCCTGAAATTTGTTTACGCTTACTGTACTGTAGTCAATTTGCTTGGAGCAACGAG 3002  
Db 3012 CTGCCCCCTGAAATTTGTTTACGCTTACTGTACTGTAGTCAATTTGCTTGGAGCAACGAG 3071  
QY 3003 CTGTGGCTGTGTACTGATCCCGAGCAATCTGCGAAAGGAAATGATAGAGGTTCTTA 3062  
Db 3072 CTGTGGCTGTGTACTGATCCCGAGCAATCTGCGAAAGGAAATGATAGAGGTTCTTA 3131  
QY 3063 TAAAGACACAGTCAAGATGCTTCCGAGCCCTCAGAGCAATTTCTATCCACGCCCCCT 3122  
Db 3132 TAAAGACACAGTCAAGATGCTTCCGAGCCCTCAGAGCAATTTCTATCCACGCCCCCT 3191  
QY 3123 GCTCAATTCAGCATGTGTGTAGAGCAGCAGATACAACTGTCTTCTTCACTCTGCTC 3182  
Db 3192 TCTGAACTCCAGCATGTGTGTAGAGCAGCAGATACAACTGTCTTCTTCACTCTGCTC 3251  
QY 3183 AGCTTCCCAATGCAACGCGCAGATGATGATCAATCAGAGCATCTGTGAGAGTGTGA 3242  
Db 3252 AGCTTCCCAATGCAACGCGCAGATGATGATCAATCAGAGCATCTGTGAGAGTGTGA 3311  
QY 3243 GAACCTGACACAGGCAACGACCTGCGAGACCTGCATATCTGCTTCTACGCTGATCCAC 3302  
Db 3312 GGACCTGACACAGGCAACGACCTGCGAGACCTGCATATCTGCTTCTATGGTGAACCCAG 3371  
QY 3303 CAATGAGGAGAAATGTGACCCATGCAAGTGCATGGGCAACGCGCATCTGTGCAACACCA 3362  
Db 3372 TAAAGAGGCAATGTGACCCATGCAAGTGCATGGGCAACGCGCATCTGTGCAACACCA 3431  
QY 3363 CACGGCAAGTGTCTTGTGACCAACGAGGCGCTCAAGGGGAGAGTGCAGCTATGTGA 3422  
Db 3432 CACTGGCAAGTGTCTTGTGACCAACGAGGCGCTCAAGGGAGAGTGCAGCTATGTGA 3491  
QY 3423 GGTAGAAATGCAATCAAGGAAACCTCTCAGAGGAAACATGTTATTTATCTTCTTAT 3482  
Db 3492 GGTAGAAATGCAATCAAGGAAACCTCTCAGAGGAAATTTGACATGTTTCAATGCTTCA 3551  
QY 3483 TGACTATCAGTTTCACTTTTGTATCCCGAGGAAACATGATGCTTATCAGAGCATCA 3542  
Db 3552 TGACTATCAGTTTCACTTTTGTATCCCGAGGAAACATGATGCTTATCAGAGCATCA 3611  
QY 3543 TTTTGTGGCTACTCTCTGACGAAACAAAACAGGATTTGGACATGTTTCAATGCTTCA 3602  
Db 3612 TTTTGTGGCTACTCTCTGATGAAACAAAACAGGATTTGGACATGTTTCAATGCTTCA 3671  
QY 3603 GAATTTCACTCAACATCACTGGGCTGCCAGTTTCTCAGTGGAAACCCAGGCTGGGA 3662  
Db 3672 AAATTTCACTCAACATCACTGGGCTGCCAGTTTCTCAGTGGAAACCCAGGCTGGGA 3731  
QY 3663 AGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAAATGA 3722  
Db 3732 AGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAAATGA 3791



```

Qy 3723 GAAGTTTGAATTTTCGCAACCAACCAATATACATCTTCTTGTGTTATGTCAGTAATTTTCC 3782
Db 3792 GAAATTTTGAATTTTCGCAACCAATATACATCTTCTTGTGTTATGTCAGTAATTTTCC 3851
Qy 3783 CTGGCCCATCAAAATTTCA 3800
Db 3852 TTGGCCCATCAAAATTTCA 3869

RESULT 12
AB093245
LOCUS AB093245 Mus musculus mRNA for mKIAA0548 protein. linear ROD 07-FEB-2003
DEFINITION AB093245 Mus musculus mRNA for mKIAA0548 protein.
ACCESSION AB093245
VERSION AB093245.1 GI:26006172
SOURCE FLI_CDNA.
ORGANISM Mus musculus (house mouse)
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Hara, Y., Nagase, T.,
Ohara, O. and Koga, H.
TITLE Prediction of the coding sequences of mouse homologues of KIAA
gene: 1. The complete nucleotide sequences of 100 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
unpublished
2 (bases 1 to 5683)
Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.
Direct Submission
Submitted (04-OCT-2002) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kanamari, Kisarazu, Chiba 292-0818, Japan
(E-mail: mouse@kazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)
The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing; Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
FEATURES
source
1..5683
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mb03041"
/tissue_type="brain"
/dev_stage="adult"
/notes="vector:modified pBC SK+"
1..5683
/gene="mKIAA0548"
<1..3181
/gene="mKIAA0548"
/notes="CDS is predicted by in silico analysis. Start codon
is not identified."
/codon_start=2
/evidence="not experimental"
/product="mKIAA0548 protein"
/protein_id="BAC41429.1"
/db_xref="GI:26006173"
/translation="AYDLTSREWLPLNHSVNSVYVYGHSLALHKDKITMYGGKIDST
GNWNELRVFIHNSVLLTPKADQAVGVGSHIVTLASGRVYMLVI FGHCPLYG
YLSVQVEYLEKNTWSILHTQCALVGGVGHSSVYDDRTKALYVHGKYAFSANKYRL
ADLRYDVDTQMTWLKDSRFRVLTHTAVYSGTMLVFGGNTHNDTSMHGAKCFSS
DFAYDIACDRWSVLPRLHDHVRPFHSVAVLYNSTMYVFGFNLSLLSDVLPVLE
QCDHRSAAACVAAGGIRCLMDTSSRCTSWELATEOEAKLSECKSEKSYRLDHRC
DHTDCYCTANTCHNDCHNCPVNSCTEGQISIAKYSCPDKNPMYCNKRTSC
RSCALDONQWEPNVECTALPENICGNWHLVGNLSCLITAKENYNDAKLSCRNHN
AFLASTISQKVEFLVQLRMLQMSQSOSKMLTLPWGLRKLINVSYWEDMSPTNS
LLQWPEPSDAGFCGLSEPSRGLKAATCINPLNGSVCPANESAKQCRTPCLAR
LACGETCSSESCECMWCSNNKQCVDSNAYVASFPQOCWEYTWSSCPPNCSGYCTCS
HCLQPGCCWCWTDPSNTGKCEGYSKGFVPMPSQASGNVYVPOPLNNSNCLSDSR
YNSWFTFCPCQNCNGKCNINOSVCEKBDLTTGKHCETCISGFGDPNGKRCQPKF
CNGHASLCLNTNTGCKFCCTTKGVGDECOLCEVENRYQGNPLKGTCTYITLLIDYQFTFS

```

## ORIGIN

```

Query Match 62.8%; Score 2693; DB 10; Length 5683;
Best Local Similarity 90.4%; Pred. No. 0;
Matches 2876; Conservative 0; Mismatches 305; Indels 0; Gaps 0;

Qy 1110 AGCGTATGACCTTGGTCTTCTAGGAGTGGCTTCCACTAAACCGTCTGTGAAACAAATGCTGT 1169
Db 1 AGCGTATGACCTTGGTCTTCTAGGAGTGGCTTCCACTAAACCGTCTGTGAAACAAATGCTGTGT 60
Qy 1170 TGTATGATATGCTCATTTCTTGGCAATATACAAGGATAAAATTTACATGATGAGGAAA 1229
Db 61 TGTATGATATGCTCATTTCTTGGCAATATACAAGGATAAAATTTACATGATGAGGAAA 120
Qy 1230 AATTGATCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTC 1289
Db 121 AATTGATCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCATATTCATTAATGAGTC 180
Qy 1290 ATGCGTGTGTTTTCACCCCTAAGCAAGGAGAGTATGAGTGGTGGGCACCTCTCCACA 1349
Db 181 ATGCGTGTGTTTTCACCCCTAAGCAAGGAGTATGAGTGGTGGGCACCTCTCCACA 240
Qy 1350 CATTGTTACACTGAAGAATGGCGAGTGGTGTATGCTGGTCACTTTGGTCACTGCCCTCT 1409
Db 241 CATTGTTACACTGGCATCTGGCGTGTGGTGTATGCTGGTCACTTTGGTCACTGCCCAT 300
Qy 1410 CTATGATATATAGCAATGTGCAAGAAATATGATTTGGATAAGAACACACATGAGTATAT 1469
Db 301 CTATGATATATAGCAATGTGCAAGAAATATGATTTGGATAAGAACACACATGAGTATAT 360
Qy 1470 ACACACCCAGGTCCTTGTGCAAGGGGTTACGGCCCATAGCAGTGTGTTACGACCATAG 1529
Db 361 ACATACCTCAGGGTCTTGTGCAAGGGGTTATGCGCACAGTAGTGTGTTATGATGACAG 420
Qy 1530 GACGAGGCGCTATAGCTTTCATGCTGGTCAAGGCTTTTCAGTGGCAATAGTACCGGCT 1589
Db 421 GACCAAGGCTCTGTAGCTTTCATGCTGGTCAAGGCTTTTCAGTGGCAATAGTACCGGCT 480
Qy 1590 TGCAAGATGATCTCTACCGATATGATGAGTATGCCAGATGAGTGGACCATCTTTAAGGACAG 1649
Db 481 TGCAAGATGATCTCTACCGATATGATGAGTATGCCAGATGAGTGGACCATCTTTAAGGACAG 540
Qy 1650 CCGATTTTCCGTTTACTTGCATACAGCTGTGTATGAGTGGAGAACCATGCTGGTGTGG 1709
Db 541 CCGATTTTCCGTTTACTTGCATACAGCTGTGTATGAGTGGAGAACCATGCTGGTGTGG 600
Qy 1710 GGGAAACACACACAAATGACATCTATGAGCCATGGCGCAATGCTTCTCTTCAGATTT 1769
Db 601 AGGAAACACACAAATGACATCTTCCATGAGCCAGGTCGCAATGCTTCTCTTCGACTT 660
Qy 1770 CATGGCCTATGACATTTGCCCTGTGACCGCTGGTCAAGTCTCCAGACCTGATCTCCACCA 1829
Db 661 CATGGCTTATGACATTTGCCCTGTGACCGCTGGTCAAGTCTCCAGACCTGATCTCCACCA 720
Qy 1830 TGATGTCAACAGATTTGGCCATTTCAGCAGTCTTACACACAGCACCATGATGTGTTGG 1889
Db 721 TGATGTCAACAGATTTGGCCATTTCAGCAGTCTTGTGTAACACAGCACCATGATGTGTTGG 780
Qy 1890 TGGTTTCAATAGTCTCTCTCCTCAGCGATCTCTGATTTTACCTTGGAGAGTGGCGATGC 1949
Db 781 CGGCTTCAACAGCTCTCTCTCCTCAGTGGAGCTTGTGTTTACCTTGGAGAGTGGCGATGC 840
Qy 1950 GCATCGGAGTGAACCGCTGTTTGTAGCAGCAGGACCTGGTATTCGGTGTGTGTCGAACAC 2009
Db 841 ACACCGAGTGAACCGCTGTTTGTAGCAGCAGGACCTGGTATTCGGTGTGTGTCGAACAC 900
Qy 2010 AGGTCGCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAAACAGAGAAAAGTATAA 2069
Db 2010 AGGTCGCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAAACAGAGAAAAGTATAA 2069

```

Db 901 ACAGTCGTCTCGATGTACCTCTCTGGAGTTGGCAACTGAAGAACAGAGAGAAAGTTAAA 960  
QY 2070 ATCAGATGTTTTCCAAAGAACTCTTGACCAATGACAGATGTGACCAACACACAGATTG 2129  
Db 961 ATCAGAGTGTGTTTTTAAAGAAACCTTGACCAATGACAGATGTGACCAACACACAGATTG 1020  
QY 2130 TTACAGCTGTACAGCCAAACACCAATGACTGCCACTGGTGCATGACCAATGTTGTCCTGAT 2189  
Db 1021 TTACAGCTGTACAGCCAAACACCAATGACTGCCACTGGTGCATGACCAATGTTGTCCTGAT 1080  
QY 2190 GAACCAAGCTGTCTAGAGCCAGATCTCCATTTTTAGGTATGAGAAATGTCCTCCAGGA 2249  
Db 1081 GAACCAAGCTGTCTAGAGCCAGATCTCCATTTTTAGGTATGAGAAATGTCCTCCAGGA 1140  
QY 2250 TAACCCCATGTACTACTGTAAACAAAGAACACAGCTGACAGAGCTGGCCCTGACACAGAA 2309  
Db 1141 TAACCCCATGTACTACTGTAAACAAAGAACACAGCTGACAGAGCTGGCCCTGACACAGAA 1200  
QY 2310 CTGCCAGTGGAGCCCGGAATCAGAGTGCATGTCCTGCGCCGAAATATCTGTGGCAT 2369  
Db 1201 CTGCCAGTGGAGCCCGGAATCAGAGTGCATGTCCTGCGCCGAAATATCTGTGGCAA 1260  
QY 2370 TGCTCGCATTTGGTTGGAACTCTCATGTTTGAATTAATCTACTGCAAGAGAAATATGA 2429  
Db 1261 TGCTCGCATTTGGTTGGAACTCTCATGTTTGAATTAATCTACTGCAAGAGAAATATGA 1320  
QY 2430 CAATGCTAAATGTTCTGTAGGAACCAAAATGCCCTTTGGCTTCTTTACAAACCAAGAA 2489  
Db 1321 CAATGCTAAATGTTCTGTAGGAACCAAAATGCCCTTTGGCTTCTTTACAAACCAAGAA 1380  
QY 2490 GAAGTGAATTTGCTTTAAGCAGCTGCAATATCTAGTCAATCTCAGAGCATGTCCAA 2549  
Db 1381 GAAGTGAATTTGCTTTAAGCAGCTGCAATATCTAGTCAATCTCAGAGCATGTCCAA 1440  
QY 2550 GCTCAGCTTAAACCCATGGTGGCTTGGAGAGTCAATGTGCTACTGCTGCTGGGA 2609  
Db 1441 GCTCAGCTTAAACCCATGGTGGTCTTGGAGAGTCAATGTGCTACTGCTGCTGGGA 1500  
QY 2610 AGATATGCTCCCATTTAAGAGTGTCTAGTGGATGCGGTCTGAGCCAGTGTGTC 2669  
Db 1501 GATATGCTCCCATTTAAGAGTGTCTAGTGGATGCGGTCTGAGCCAGTGTGTC 1560  
QY 2670 TGATTTCTGTGAAATTTTACAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2729  
Db 1561 TGCTTTCTGTGGATCTTGTAGAGCTTGTACTCGGGGATTAAGGCTGCAACCTGCAT 1620  
QY 2730 CAACCCATCAATGTAGTGTGTGAAAGGCTTGCACCAACACAGTGTCTAGAGTGGCG 2789  
Db 1621 CAACCCATCAATGTAGTGTGTGAAAGGCTTGCACCAACACAGTGTCTAGAGTGGCG 1680  
QY 2790 GACACATGTGCTTGGAGCAGCATGTGGAGATGTGACAGCGGAGCTCTGAGTGCAT 2849  
Db 1681 GACACATGTGCTTGGAGCAGCATGTGGAGATGTGACAGCGGAGCTCTGAGTGCAT 1740  
QY 2850 GTGGTGCAGCAACATGAAGCAGTGTGTGAGTCTCAATGCTTATGTGGCTCTTCCCTTT 2909  
Db 1741 GTGGTGCAGCAACATGAAGCAGTGTGTGAGTCTCAATGCTTATGTGGCTCTTCCCTTT 1800  
QY 2910 TGGCCAGTGTATGNAATGTTATACATGAGCAGCATGCTGCCCCCTGAAATTTGTTAGGCTA 2969  
Db 1801 TGGCCAGTGTATGNAATGTTATACATGAGCAGCATGCTGCCCCCTGAAATTTGCTTGGCTA 1860  
QY 2970 CTGTACCTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGTGTACTGATCTCCAGCAA 3029  
Db 1861 TTGTACCTGTGAGCAATGCTTGGAGCAGCAGGCTGTGGTGTGTACTGATCTTAGCAA 1920  
QY 3030 TACTGGCAAGAGGAAATGCTATAGAGGTTCTCTATAAGAGCAGTGAAGATGCTTGCCTA 3089  
Db 1921 TACTGGCAAGAGGAAATGCTATAGAGGTTCTCTATAAGAGCAGTGAAGATGCTTGCCTA 1980  
QY 3090 AGCCCTCTACAGAAATTTCTATCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGGA 3149  
Db 1981 GGCCTCTGACAGAAATGTTATCCACAGCCCTTCTGAACTCCAGCATGTGTCTAGAGGA 2040

QY 3150 CAGCAGATACAACTGGTCTTTTCATTTCACTGTCCAGCTTGCATATGCAACGGCCACAGTAA 3209  
Db 2041 CAGCAGATACAACTGGTCTTTTCATTTCACTGTCCAGCTTGCATATGCAACGGCCACAGTAA 2100  
QY 3210 ATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACTGACCAACAGGCAAGCACTGCGA 3269  
Db 2101 ATGCATCAACAGAGTGTCTGTGAGAAGTGTGAGAACTGACCAACAGGCAAGCACTGCGA 2160  
QY 3270 GACCTGCATATCTGGCTTTCTAGCGGTATCCCAACATGAGGAGAAATGTGAGCCATGCAA 3329  
Db 2161 GACCTGCATATCTGGCTTTCTATGTGACCCGACTTAATGAGGAGAAATGTGAGCCATGCAA 2220  
QY 3330 GTGCAATGGGCAAGCGCTCTGTGCAACCAACAGGCAAGTGTCTGCAACCAACAA 3389  
Db 2221 GTGCAATGGGCAAGCGCTCTGTGCAACCAACAGGCAAGTGTCTGCAACCAACAA 2280  
QY 3390 GGGCTCAAGGGGAGCAGTGCAGCTATGTAGGTAGAAATCGATACCAAGAAACCC 3449  
Db 2281 AGGTGTCAAGGGGAGCAGTGCAGCTATGTAGGTAGAAATCGATACCAAGAAATCC 2340  
QY 3450 TCTCAGAGGAACATGTTATATATCTTCTTATTGACTATCAGTTCACTTTAGTCTATC 3509  
Db 2341 TCTCAGAGGAACATGTTATATATCTTCTTATTGACTATCAGTTCACTTTAGTCTATC 2400  
QY 3510 CCAGGAAGATGATCGCTTTATACAGCATATCAATTTTGGCTTACTCTCGACCAACAAA 3569  
Db 2401 CCAGGAAGATGATCGCTTTATACAGCATATCAATTTTGGCTTACTCTCGACCAACAAA 2460  
QY 3570 CAGGATTTGGACATGTTTATCATCATGCTCCAAAGAAATTTCAACCTCAACATCACTGGGC 3629  
Db 2461 CAGGATTTGGACATGTTTATCATCATGCTCCAAAGAAATTTCAACCTCAACATCACTGGGC 2520  
QY 3630 TGCCAGTTTCTCAGCTGGAACCCAGCTGAGAGAGATGCTGTGTTTCAAAACCAA 3689  
Db 2521 CACAGCTTTCCAGCCGGAACCCAGCTGAGAGAGATGCTGTGTTTCAAAACCAA 2580  
QY 3690 CATTAAAGAGTACAAAGATAGTTTCTTAATGAGAAGTTTGAATTTGCAACCAACCAA 3749  
Db 2581 CATTAAAGAGTACAAAGATAGTTTCTTAATGAGAAGTTTGAATTTGCAACCAACCAA 2640  
QY 3750 TATCAGCTTTCTTGTGTTATGTCAGTAATTTCACTGGCCCAATCAAAATTCAGATGCTT 3809  
Db 2641 CATCAGCTTTCTTGTGTTATGTCAGTAATTTCACTGGCCCAATCAAAATTCAGATGCTT 2700  
QY 3810 CTCTCAGCAGCAATTTTATGGAACCTGCTACAGTTCTTCTGTCAGTTCTTCTCAGTTGTTT 3869  
Db 2701 CTCTCAGCAGCAATTTTATGGAACCTGCTACAGTTCTTCTGTCAGTTCTTCTCAGTTGTTT 2760  
QY 3870 CCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3929  
Db 2761 CCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820  
QY 3930 CAGACGTAGAGAGCAACTTTCTTCGAGAGATGCAACAGATGGCCAGCCGCTTTCCTCTC 3989  
Db 2821 CAGGCGGAGAGAGCAACTTTCTTCGAGAGATGCAACAGATGGCCAGCCGCTTTCCTCTC 2880  
QY 3990 TGTAAATGTGCTTGGAAACAGATGAGAGCTCCTGATCTTATTTGGGGGAGTATATA 4049  
Db 2881 TGTAAACGTGTGCTTGGAAACAGATGAGAGCTCCTGATCTCTATTTGGGGGAGTATATA 2940  
QY 4050 GACTGTTTCCAAACCATTTGCACTGGAGCCGTTTGGCAACAAAGCCGCTGCTCTCTC 4109  
Db 2941 GACCGTTCCTAGACCCATTTGCCCTGGAGCCCTGCTTTGGTAAAGCCGAGTCTCTCTC 3000  
QY 4110 TGTGTTGTGAGCTTCTCTGAGGCTGGTGGCATCCCTCTCTCTGGCAGTCAAGTCT 4169  
Db 3001 TGTATTTCTGAGGCTTCTCTGAGGCTGGGAGAAATCCCTCTCTCTGCTGCTGCTGCTGCT 3060  
QY 4170 TCTGTGGCCAGGCTTCTGAGCAATTTCTCAGCAGATGCTGAGTGTACAAAGGAGAA 4229  
Db 3061 CCGTGTGGCCAGTGGCTTCTGAGCAATTTCTCAGCAGATGCTGAGTGTACAAAGGAGAA 3120



QY 1683 AGTGAGTGGAAACCATGCTGGTGTGGGGGAAACACACACAATGACATCTATGAGCCA 1742  
Db |||||  
QY 1680 AGTGAGTGGAAACCATGCTGGTGTGGGGGAAACACACACAATGACATCTATGAGCCA 1739  
Db |||||  
QY 1743 TGGCGGCAAAATGCTCTCTTCAGATTTTCATGGCTATGACATTTGCTGTGACCGCTGGTC 1802  
Db |||||  
QY 1740 TGGCGGCAAAATGCTCTCTTCAGATTTTCATGGCTATGACATTTGCTGTGACCGCTGGTC 1799  
Db |||||  
QY 1803 AGTGCTTCCGAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1862  
Db |||||  
QY 1800 AGTGCTTCCGAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1859  
Db |||||  
QY 1863 ACACAAACAGCACCATGATGTGTGGTGGTTCATAGTCTCCTCCCTCAGCGACATCCT 1922  
Db |||||  
QY 1860 ACACAAACAGCACCATGATGTGTGGTGGTTCATAGTCTCCTCCCTCAGCGACATCCT 1919  
Db |||||  
QY 1923 GGTATTCACTCGGAACAGTGTGATCGGCATCGGAGTGAAGCGCTTTGTTAGCAGCAGG 1982  
Db |||||  
QY 1920 GGTATTCACTCGGAACAGTGTGATCGGCATCGGAGTGAAGCGCTTTGTTAGCAGCAGG 1979  
Db |||||  
QY 1983 ACTGTGATTTCGGTGTGTGGAAACACACAGGTCCTCAGTGTATCTCGTGGCGCTGGC 2042  
Db |||||  
QY 1980 ACTGTGATTTCGGTGTGTGGAAACACACAGGTCCTCAGTGTATCTCGTGGCGCTGGC 2039  
Db |||||  
QY 2043 AACTGATGAACAAAGAAAGTTAAATTCAGATTTTTCCTCAAAAGAACTCTTGACCA 2102  
Db |||||  
QY 2040 AACTGATGAACAAAGAAAGTTAAATTCAGATTTTTCCTCAAAAGAACTCTTGACCA 2099  
Db |||||  
QY 2103 TGACAGATGTGACCAAGCAACAGATTTTACAGTGTGACAGCCCAACCAATGATGACCA 2162  
Db |||||  
QY 2100 TGACAGATGTGACCAAGCAACAGATTTTACAGTGTGACAGCCCAACCAATGATGACCA 2159  
Db |||||  
QY 2163 CTGGTGCATGACCAATGTGTCCCAAGCAACACAGTGTCTCAGAGCCAGATCTCCAT 2222  
Db |||||  
QY 2160 CTGGTGCATGACCAATGTGTCCCAAGCAACACAGTGTCTCAGAGCCAGATCTCCAT 2219  
Db |||||  
QY 2223 TTTTAGGTATGAGATTTGCCCAAGGATAACCCATGATCTACTGTAAACAAAGAGACAG 2282  
Db |||||  
QY 2220 TTTTAGGTATGAGATTTGCCCAAGGATAACCCATGATCTACTGTAAACAAAGAGACAG 2279  
Db |||||  
QY 2283 CTGAGGAGCTGTGCTCCGACCAAGCACTGCTGAGGAGCCCGGAAATCAGGAGTGCAT 2342  
Db |||||  
QY 2280 CTGAGGAGCTGTGCTCCGACCAAGCACTGCTGAGGAGCCCGGAAATCAGGAGTGCAT 2339  
Db |||||  
QY 2343 TGCCCTGCCGAAATATCTGTGGCAATGGCTGGCAATTTGGTTGG 2387  
Db |||||  
QY 2340 TGCCCTGCCGAGTGGCTTCAGGGTCTATCTTGGTGTGTGGG 2384  
Db |||||

## RESULT 14

AR164815  
LOCUS AR164815 6370 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 12 from patent US 6274339.  
ACCESSION AR164815  
VERSION AR164815.1 GI:16238024  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
AUTHORS 1 (bases 1 to 6370)  
TITLE Moore,K. and Nagle,D.Lynn.  
METHODS Methods and compositions for the diagnosis and treatment of body  
weight disorders, including obesity  
JOURNAL Patent: US 6274339-A 12-14-AUG-2001;  
FEATURES location/Qualifiers  
1. 6370  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 43.2%; Score 1853; DB 6; Length 6370;  
Best Local Similarity 96.3%; Pred. No. 0;  
Matches 1980; Conservative 7; Mismatches 47; Indels 23; Gaps 8;

QY 2257 ATGTACTACTGTAAACAAGAACAGCTGTCAGGAGCTGTGCCCTGGACCAAGTCTCCAG 2316  
Db |||||  
QY 2317 TGGAGCCCGGAAATCAGGAGTGCATTTGCCCTGCCGAAAAATATCTGTGGCATTTGGCTGG 2376  
Db |||||  
QY 61 TGGAGCCCGGAAATCAGGAGTGCATTTGCCCTGCCGAAAAATATCTGTGGCATTTGGCTGG 120  
Db |||||  
QY 2377 CATTTGGTTGGAAACTCATGTTTGAATAATTTACTACTGCCAAGGAGAAATATGCAATGCT 2436  
Db |||||  
QY 121 CATTTGGTTGGAAACTCATGTTTGAATAATTTACTACTGCCAAGGAGAAATATGCAATGCT 180  
Db |||||  
QY 2437 AAATTGTTCTGTAGGAACACAAATGCTTTTGGCTTCTTTACAAACCAGAAAGGTA 2496  
Db |||||  
QY 181 AAATTGTTCTGTAGGAACACAAATGCTTTTGGCTTCTTTACAAACCAGAAAGGTA 240  
Db |||||  
QY 2497 GAATTTGCTTAAAGCAGCTGCGAATAATGCAATGCTTCTCAGAGCATGTCCAAAGTCTACC 2556  
Db |||||  
QY 241 GAATTTGCTTAAAGCAGCTGCGAATAATGCAATGCTTCTCAGAGCATGTCCAAAGTCTACC 300  
Db |||||  
QY 2557 TTAACCCCATGGGTGCGCCTT---CGAAAGATCAATGTCTCTACTGG---TGTGGG 2608  
Db |||||  
QY 301 TTAACCCCATGGGTGCGCCTTTCGGGAAGTTCATGTCTCTACTTGGTCTGGGGAAG 360  
Db |||||  
QY 2609 AAGATATGTCCTTAAAGCAGCTGCGAATAATGCAATGCTTCTCAGAGCATGTCCAAAGTCTACC 2663  
Db |||||  
QY 361 GATATGTCCTTAAAGCAGCTGCGAATAATGCAATGCTTCTCAGAGCATGTCCAAAGTCTACC 420  
Db |||||  
QY 2664 TGATGCT---GGATTCCTGTGGAAATTTT---ATCAGAACCCAGTACTCGGGGACTGAAGG 2716  
Db |||||  
QY 421 TGTGCTTGGTATCTGTGGAAATTTTATTCAGGAACCCAGTACTTCGGGACTGAAGG 480  
Db |||||  
QY 2717 CTGCAACCTGCAT---CAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACACAG 2774  
Db |||||  
QY 481 CTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACACAG 540  
Db |||||  
QY 2775 TGCTAA---GCAGTGGCGGACACCATGTGCTTCAGACAGCATGTGGAGATTCACACAGG 2833  
Db |||||  
QY 541 TGCTAAGGAGTGGCGGACACCATGTGCTTCAGACAGCATGTGGAGATTCACACAGG 600  
Db |||||  
QY 2834 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGAGCTTCCAAATGCTATG 2893  
Db |||||  
QY 601 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGAGCTTCCAAATGCTATG 660  
Db |||||  
QY 2894 TGGCCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTTATACGATGAGCACTGCCCCCTG 2953  
Db |||||  
QY 661 TGGCCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTTATACGATGAGCACTGCCCCCTG 720  
Db |||||  
QY 2954 AAAATTGTTGAGGCTACTGTACTGTAGTCTATGCTTGGAGCAACAGGCTGTGCTGCT 3013  
Db |||||  
QY 721 AAAATTGTTGAGGCTACTGTACTGTAGTCTATGCTTGGAGCAACAGGCTGTGCTGCT 780  
Db |||||  
QY 3014 GTACTGTATCCAGCAATACTGCGAAGGGAATGCAATAGAGGTTCTTATTAAGGACAG 3073  
Db |||||  
QY 781 GTACTGTATCCAGCAATACTGCGAAGGGAATGCAATAGAGGTTCTTATTAAGGACAG 840  
Db |||||  
QY 3074 TGAAGATGCTTTGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTCCA 3133  
Db |||||  
QY 841 TGAAGATGCTTTGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTCCA 900  
Db |||||  
QY 3134 GCATGTGTCTAGAGCAGACAGATACAACTGGTCTTTTCAITTCATGTCCAGCTTGCAT 3193  
Db |||||  
QY 901 GCATGTGTCTAGAGCAGACAGATACAACTGGTCTTTTCAITTCATGTCCAGCTTGCAT 960  
Db |||||  
QY 3194 GCAACGGCCACAGTAAATGCAATCAATCAGACATCTGTGAGAGTGTGAGAACCTGACCA 3253  
Db |||||  
QY 961 GCAACGGCCACAGTAAATGCAATCAATCAGACATCTGTGAGAGTGTGAGAACCTGACCA 1020  
Db |||||  
QY 3254 CAGGCAAGCACTGGAGACCTGCAATATCTGGCTTCTTACGGTGTATCCCAACCAATGGAGGA 3313  
Db |||||  
QY 1021 CAGGCAAGCACTGGAGACCTGCAATATCTGGCTTCTTACGGTGTATCCCAACCAATGGAGGA 1080  
Db |||||





Result No.	Score	Query Match	Length	DB ID	Description
1	3429.2	79.9	3976	29	AY118587 Homo sapi
2	2923.6	68.1	3976	29	AY118589 Mus muscu
3	2864.4	66.8	3971	29	AY118588 Pan trogl
4	925.6	21.6	1201	13	BX440935 BX440935

[illegible]

AY418587  
LOCUS  
AY418587 3976 bp DNA linear  
VIRTUAL TRANSCRIPT. partial sequence.  
GSS 17-DEC-2003



JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.

FEATURES Location/Qualifiers

source 1..3976

organism="Homo sapiens"

mol\_type="genomic DNA"

db\_xref="taxon:9606"

gene <1..>3976

gene="ATRN"

locus\_tag="HCM6607"

ORIGIN

Query Match 79.9%; Score 3429.2; DB 29; Length 3976;

Best Local Similarity 86.3%; Pred. No. 0;

Matches 3431; Conservative 0; Mismatches 545; Indels 0; Gaps 0;

QY 315 TGTCAACGCGCTGCTGCAACCCCTGGCACCGCCAGTGGCTCTGCCCGCGCGGTGGT 374

Db 1 TGTCAACGCGCTGCTGCAACCCCTGGCACCGCCAGTGGCTCTGCCCGCGCGGTGGT 60

QY 375 GGGCGAGCAATCCAGCACTGCGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTGGT 434

Db 61 GGGCGAGCAATCCAGCACTGCGGGGGCGGCTTCAGNNNNNNNNNNNNNNNNNNNN 120

QY 435 GACAGATGACCTGGAAATATATAAATAAACAAGAGTGGACGTGGCTCAATGAAGACA 494

Db 121 NNN 180

QY 495 GCCAATAGAAATATAGACTTCGTTTCAATTTGCTACAGAGTGTAGTGGACCA 554

Db 181 GCCAATAGAAATATAGACTTCGTTTCAATTTGCTACAGAGTGTAGTGGACCA 240

QY 555 TTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTCTGCATTTAGTGGCCT 614

Db 241 TTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTCTGCATTTAGTGGCCT 300

QY 615 CATTGTTCTTGAGAGATGCAATGAGACTGTCCCTGAGGTGTTGCCACATCAGTTA 674

Db 301 NNN 360

QY 675 TGCCTTGCTGCAATTTTCTAGTGTGCTTATAATTTGACTGGAATTAATTAATTA 734

Db 361 NNN 420

QY 735 CAGTTTGTATGTTGTCAAATAACTGCTCAGGCCGAGAGAGTGAAGTACAGTAATAG 794

Db 421 NNNTTTGTATGTTGTCAAATAACTGCTCAGGCCGAGAGAGTGAAGTACAGTAATAG 480

QY 795 CAGCGAACTCTTCAATGTGAATGTTCTGAAATCTGAAAGTGAAGCATGTGACATTC 854

Db 481 CAGCGAACTCTTGAATGTGAATGTTCTGAAATCTGAAAGTGAAGCATGTGACATTC 540

QY 855 TCACGTACAGAACTGTGGTTTCTCATCGAGGCAATCTGCAATTCAAAGTCAATGTCAG 914

Db 541 TCACGTACAGAACTGTGGTTTCTCATCGAGGCAATCTGCAATTCAAAGTCAATGTCAG 600

QY 915 AGGATGCTCCTGCTTCAGACTGCGAGGTCTGAGTGTTCAGTCTGCTGACCTGACCTAA 974

Db 601 AGGATGCTCCTGCTTCAGACTGCGAGGTCTGAGTGTTCAGTCTGCTGACCTGACCTAA 660

QY 975 CCAGTCAATTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCGAGAGCATCTCATAA 1034

Db 661 CCAGTCAATTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCGAGAGCATCTCATAA 720

QY 1035 AGCTGTGCTCAATGGAACATATATGGGTGTTGAGGATATATGTTCAACACTCAGA 1094

Db 721 AGCTGTGCTCAATGGAACATATATGGGTGTTGAGGATATATGTTCAACACTCAGA 780

QY 1095 TTATAACATGGTTCTAGCGTATGACCTTGTCTAGGGAGTGGCTTCCACTAACCGTTTC 1154

Db 781 TTATAACATGGTTCTAGCGTATGACCTTGTCTAGGGAGTGGCTTCCACTAACCGTTTC 840

QY 1155 TGTCAACAAATGCTGTTGTAGATATGCTCATCTTTGGCATATATACAAGGATAAAATTTA 1214

Db 841 TGTCAACAAATGCTGTTGTAGATATGCTCATCTTTGGCATATATACAAGGATAAAATTTA 900

QY 1215 CATGTATGGAGGAAAAATTCATCAACTGGAAATGTGACCAATGCTGAGAGTTTTC 1274

Db 901 NNN 960

QY 1275 CATTCAATAGACTCATGGGTGTTGTGACCCCTAAGGCAAGGAGCAGTATGCGAGTGT 1334

Db 961 NNN 1020

QY 1335 TGGGCACTCTGCACACATTTGTACACTGAAGAATGCCGAGTGGTCACTGCTGCTCATCTT 1394

Db 1021 NNN 1080

QY 1395 TGTCTACTGCTCTCTATGGATATATAAGCAATGTGCAAGAAATATGATTTGGATAAGAA 1454

Db 1081 NNN 1140

QY 1455 CACATCGAGTATATACACACCCAGGTCCTCTGTGCAAGGGGTACGGCCATAGCAG 1514

Db 1141 CACATCGAGTATATACACACCCAGGTCCTCTGTGCAAGGGGTACGGCCATAGCAG 1200

QY 1515 TGTTTACGACCATAGGACCAAGGCGCTTATACGTTTATGTTGGCTTACAAGGCTTTCAGTGC 1574

Db 1201 TGTTTACGACCATAGGACCAAGGCGCTTATACGTTTATGTTGGCTTACAAGGCTTTCAGTGC 1260

QY 1575 CAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGCGATATGCCAGATGTGAC 1634

Db 1261 CAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGCGATATGCCAGATGTGAC 1320

QY 1635 CATTTCTAAGGACAGCCGATTTTCCGTTACTTTCACACAGCTGTGATAGTGAAGTGAAC 1694

Db 1321 CATTTCTAAGGACAGCCGATTTTCCGTTACTTTCACACAGCTGTGATAGTGAAGTGAAC 1380

QY 1695 CATGCTGTTGTTGGGGGAAACACACAAATGACATCTATAGCCCATGCGCCAAATG 1754

Db 1381 CATGCTGTTGTTGGGGGAAACACACAAATGACATCTATAGCCCATGCGCCAAATG 1440

QY 1755 CTTCTCTTCAGATTTTCAATGCTTATGACCTGCTGACCGCTGCTGCTGCTGCTGCTGCTGCT 1814

Db 1441 CTTCTCTTCAGATTTTCAATGCTTATGACCTGCTGACCGCTGCTGCTGCTGCTGCTGCTGCT 1500

QY 1815 ACTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGCAGTCTTACACAGCAGCAC 1874

Db 1501 NNN 1560

QY 1875 CATGTATGTTGCTGCTGCTTCAATGATCTCTCTCAGCGACATCTGCTGATTCACCTC 1934

Db 1561 CATGTATGTTGCTGCTGCTTCAATGATCTCTCTCAGCGACATCTGCTGATTCACCTC 1620

QY 1935 GGACAGTGTGATGCGCATCGAGTGAAGCGCTGTTGTTAGCAGCAGACCTGCTGATTCG 1994

Db 1621 GGACAGTGTGATGCGCATCGAGTGAAGCGCTGTTGTTAGCAGCAGACCTGCTGATTCG 1680

QY 1995 GGTGCTGTGGAACACAGGCTGCTCAGTGTATCTGCTGGGCGCTGGCACTGATGAACA 2054

Db 1681 GGTGCTGTGGAACACAGGCTGCTCAGTGTATCTGCTGGGCGCTGGCACTGATGAACA 1740

QY 2055 AGAAGAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTCTTCACCATGACAGATGTGA 2114

Db 1741 AGAAGAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTCTTCACCATGACAGATGTGA 1800

QY 2115 CCAGCACAACAGATTTGTTACAGCTGTACAGCACAACCAATGACTGCTGCTGCTGCTGCTGCT 2174

Db 1801 CCAGCACAACAGATTTGTTACAGCTGTACAGCACAACCAATGACTGCTGCTGCTGCTGCTGCT 1860

QY 2175 CCATTTGCTGCCAGGAAACACAGCTGCTCAGAGGCGCAGATCTCCATTTTGGTATGA 2234

Db 1861 CCATTTGCTGCCAGGAAACACAGCTGCTCAGAGGCGCAGATCTCCATTTTGGTATGA 1920



QY	2235	GAATTGCCCCAAGATTAACCCCATGTACTACTGTAAACAAGACACAGCTGCGAGGAGCTG	2294
Db	1921	GAATTGCCCCAAGATTAACCCCATGTACTACTGTAAACAAGACACAGCTGCGAGGAGCTG	1980
QY	2295	TGCCCTGGACAGCAAGCTGCGAGTGGAGCCCGGAATCAGAGTGCAATTCGCCCTGCCCGA	2354
Db	1981	TGCCCTGGACAGCAAGCTGCGAGTGGAGCCCGGAATCAGAGTGCAATTCGCCCTGCCCGA	2040
QY	2355	AAATATCTGTGGCAATGGCTGGCAATTTGGTGGAAATCTCATTTTGAATAATTAAGTATGC	2414
Db	2041	AAATATCTGTGGCAATGGCTGGCAATTTGGTGGAAATCTCATTTTGAATAATTAAGTATGC	2100
QY	2415	CAAGGAAATTAATCAGCAATGCTAAATTTCTGTAGGAACACCAATCCCTTTTGGCTTC	2474
Db	2101	CAAGGAAATTAATCAGCAATGCTAAATTTCTGTAGGAACACCAATCCCTTTTGGCTTC	2160
QY	2475	TCTTACACCCAGAGAGTGAATTTGTCCTTTAAGCAGCTGCGAATAATGCAAGTCAATC	2534
Db	2161	TCTTACACCCAGAGAGTGAATTTGTCCTTTAAGCAGCTGCGAATAATGCAAGTCAATC	2220
QY	2535	TCAGAGCATGTCAGCTCACTTAACCCCATGGTGGGCTTCGGAGATCAATGTGTC	2594
Db	2221	TCAGAGCATGTCAGCTCACTTAACCCCATGGTGGGCTTCGGAGATCAATGTGTC	2280
QY	2595	CTACTGGTGGGAGATATGTCCTCCCATTTTACAAATAGATTTTACTACAGTGGATGCCGTC	2654
Db	2281	CTACTGGTGGGAGATATGTCCTCCCATTTTACAAATAGATTTTACTACAGTGGATGCCGTC	2340
QY	2655	TGAGCCAGTGAATGCTGGAATTTTATCAGAACCCAGTACTGGGAGCTGAA	2714
Db	2341	TGAGCCAGTGAATGCTGGAATTTTATCAGAACCCAGTACTGGGAGCTGAA	2400
QY	2715	GGCTGCACTGCATCAACCCCACTCAATGGTGTGTCTGTAAGGCTGCAAAACCAACAG	2774
Db	2401	GGCTGCACTGCATCAACCCCACTCAATGGTGTGTCTGTAAGGCTGCAAAACCAACAG	2460
QY	2775	TGCTAAGCAGTCCCGGACCACTATGCTTCCAGACAGCATGTGGAGATTCACACAGCGG	2834
Db	2461	TGCTAAGCAGTCCCGGACCACTATGCTTCCAGACAGCATGTGGAGATTCACACAGCGG	2520
QY	2835	CAGCTCTAGTGCATGTGGTGGAGCAACATGAGCAGTGTGGACTCCCAATGCTATGT	2894
Db	2521	CAGCTCTAGTGCATGTGGTGGAGCAACATGAGCAGTGTGGACTCCCAATGCTATGT	2580
QY	2895	GGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCTCCCTGA	2954
Db	2581	GGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCTCCCTGA	2640
QY	2955	AAATTTGTCAGGCTACTGCTAGTATGCTTGGAGCAACAGAGCTGTGGCTGTG	3014
Db	2641	AAATTTGTCAGGCTACTGCTAGTATGCTTGGAGCAACAGAGCTGTGGCTGTG	2700
QY	3015	TACTGATCCAGCAATACCTGGCAAGGAAATGCAATAGAGGTTTCTTATAAGGACCACT	3074
Db	2701	TACTGATCCAGCAATACCTGGCAAGGAAATGCAATAGAGGTTTCTTATAAGGACCACT	2760
QY	3075	GAAGATGCTTCCGAGCCCTTACAGGAAATTTCTATCCAGAGCCCTGCTCAATTCAG	3134
Db	2761	GAAGATGCTTCCGAGCCCTTACAGGAAATTTCTATCCAGAGCCCTGCTCAATTCAG	2820
QY	3135	CATGTGCTTAGAGCAGCAGATACAACTGGTCTTCTTCTACTCTCCAGCTTGGCAATG	3194
Db	2821	CATGTGCTTAGAGCAGCAGATACAACTGGTCTTCTTCTACTCTCCAGCTTGGCAATG	2880
QY	3195	CAAGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGACCAAC	3254
Db	2881	CAAGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGACCAAC	2940
QY	3255	AGGCAAGCATCTGGAGACCTGATCTGGCTTCTACGGTGTATCCCAACCAATGGAGGAA	3314
Db	2941	AGGCAAGCATCTGGAGACCTGATCTGGCTTCTACGGTGTATCCCAACCAATGGAGGAA	3000
QY	3315	ATGTACGCCATGCAAGTGGCAACGCGTCTCTGTGCAACACCAACACAGCGGCAAGTG	3374

Db	3001	ATGTACGCCATGCAAGTGCATGGGCAACGCGTCTCTGTGCAACACCAACACGGGCAAGTG	3060
QY	3375	CTTCTGCAACCAACAAAGGGGTCAGAGGGGACAGATGCGAGCTATGTGAGTGAAGAAATCG	3434
Db	3061	CTTCTGCAACCAACAAAGGGGTCAGAGGGGACAGATGCGAGCTATGTGAGTGAAGAAATCG	3120
QY	3435	ATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTATATCTCTTCTTATTGACTATCAGTT	3494
Db	3121	ATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTATATCTCTTCTTATTGACTATCAGTT	3180
QY	3495	CACCTTTAGTCTATCCAGGAAAGATGATGCTATTATACAGCTATCAATTTTGGCTTAC	3554
Db	3181	CACCTTTAGTCTATCCAGGAAAGATGATGCTATTATACAGCTATCAATTTTGGCTTAC	3240
QY	3555	TCTGAGCAACCAAAACAGGGATTTGGACATGTTCAATGCTCCCAAGAAATTTCAACCT	3614
Db	3241	TCTGAGCAACCAAAACAGGGATTTGGACATGTTCAATGCTCCCAAGAAATTTCAACCT	3300
QY	3615	CAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTCT	3674
Db	3301	CAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTCT	3360
QY	3675	TGTTTCAAAAACCAACATTAAGAGTACAAAGATGTTTCTCTAATGAGAAATTTGATTT	3734
Db	3361	TGTTTCAAAAACCAACATTAAGAGTACAAAGATGTTTCTCTAATGAGAAATTTGATTT	3420
QY	3735	TGCAACCAACCAAAATATCACTTTCTTTTATGTCAGTAAATTTTCACTGGCCCACTAA	3794
Db	3421	TGCAACCAACCAAAATATCACTTTCTTTTATGTCAGTAAATTTTCACTGGCCCACTAA	3480
QY	3795	AATTCAGATGCTCTCTCAGCAGCAAGCAATTTTATGAGCCTGGTACAGTTCTTCTGAC	3854
Db	3481	AATTCAGATGCTCTCTCAGCAGCAAGCAATTTTATGAGCCTGGTACAGTTCTTCTGAC	3540
QY	3855	TTTCTTTCAGTTGTTCTCTCTCTTTGCTCTGCTGGTGGCTGCTGTTTGGAAAGATCAACA	3914
Db	3541	TTTCTTTCAGTTGTTCTCTCTCTTTGCTCTGCTGGTGGCTGCTGTTTGGAAAGATCAACA	3600
QY	3915	AAGTTGTGGGCTCCAGAGCTAGAGAGCAACCTTTTCAGAGATGCAACAGATGGCCAG	3974
Db	3601	AAGTTGTGGGCTCCAGAGCTAGAGAGCAACCTTTTCAGAGATGCAACAGATGGCCAG	3660
QY	3975	CCGTCCCTTTGCTCTGTAAATGTCGCCCTTGGAAACAGATGAGAGCCTCTCTGATCTTAT	4034
Db	3661	CCGTCCCTTTGCTCTGTAAATGTCGCCCTTGGAAACAGATGAGAGCCTCTCTGATCTTAT	3720
QY	4035	TGGGGGAGTATAAAGACTGTTCCCAACCCATTGCACTGGAGCCCTGTTTGGCAACA	4094
Db	3721	TGGGGGAGTATAAAGACTGTTCCCAACCCATTGCACTGGAGCCCTGTTTGGCAACA	3780
QY	4095	AGCCGCTGCTCTCTGTGTTTGTAGGCTCCCTCGAGCCTGGGTGGCATCCCTCTCC	4154
Db	3781	AGCCGCTGCTCTCTGTGTTTGTAGGCTCCCTCGAGCCTGGGTGGCATCCCTCTCC	3840
QY	4155	TGGGCACTCAGGCTTCTGCTGCGCAGCGCTGGTGGACATTTCTCAGCAGATGCCGAT	4214
Db	3841	TGGGCACTCAGGCTTCTGCTGCGCAGCGCTGGTGGACATTTCTCAGCAGATGCCGAT	3900
QY	4215	AGTGTACAGGAGAGTCAAGAGCCGCTGAGAAACCGGAGCAGCAGCCCTCTGACAGCC	4274
Db	3901	AGTGTACAGGAGAGTCAAGAGCCGCTGAGAAACCGGAGCAGCAGCCCTCTGACAGCC	3960
QY	4275	TGGGACCTGCATCTGA 4290	
Db	3961	TGGGACCTGCATCTGA 3976	

RESULT 2  
AY418589  
LOCUS  
DEFINITION Mus musculus ATN gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.

3976 bp DNA linear GSS 17-DEC-2003



3015	QY	TACTGATCCAGCAATAC	TGCGAAGGGAATGCATAGAGGTTCTCTATTAAGAGCACG	3074
2701	DB	TACTGATCTTAGCAATAC	TGCGAAGGGAATGTATTGAGGGCAGCTATTAAGAGCACTGT	2760
3075	QY	GAAGATGCGTTTGGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTC	CAAG	3134
2761	DB	GAAGATGCGGTACAGGCCCTCTG	CAGGAAATGTGTATCCACAGCCCTTCTGAATCCAG	2820
3135	QY	CATGTGTCTAGAGGACAGCAGATACAACTGTGCTTTCA	TTCACCTGTGCCACTTGGCCATG	3194
2821	DB	CATGTGTCTAGAGGACAGCAGATACAACTGTGCTTTCA	TTCACCTGTGCCACTTGGCCATG	2880
3195	QY	CAACGGCCACAGTAATATGCATCAATCAGACGATCTGTGAGAAAGTGTGAGAACTG	GAGCAC	3254
2881	DB	CAACGACACACGAATATGCATCAACAGAGTATCTGTGAGAAAGTGTGAGGACCTG	GACCA	2940
3255	QY	AGCAAGACACTGAGGACCTGCTCATATCTGGCTTCTTA	CGGTATCCCAACCAATGAGGGAA	3314
2941	DB	GGGCAAGCACTGCGAGACCTGCTATCTGGCTTCTAT	TGGTACCCGCTAATGGAGGCAA	3000
3315	QY	ATGTGAGCCATGCAAGTGTCAATGGCAGCGTCTCTGTGCAACACCAACGCGGCAAGT	G	3374
3001	DB	ATGTGAGCCATGCAAGTGTCAATGGCAGCGATCACTGTGCAACACCAACCGCGCAAGT	G	3060
3375	QY	CTTCTGCACCAACGAGGGCGTCAAGGGGAGCAGTGCAGCTATGTGAGGTAGAAAAATCG		3434
3061	DB	CTTCTGTACCAACAAAGTGTCAAGGGGAGCGAGTGCAGCTATGTGAGGTAGAAAAATCG		3120
3435	QY	ATACCAAGGAACCTCTCAGAGGAACATGTTATATATCTCTTATTCGACTATCAGTT		3494
3121	DB	ATACCAAGGAACCTCTCTCAAGGGAACATGCTATCTATACCCTTCTCTATTCGACTATCAGTT		3180
3495	QY	CACCTTTAGTCTATCCCAAGGAAGATGATCGCTATTATACACAGCTATCAATTTTGTGGCTAC		3554
3181	DB	CACCTTTAGCTCTCCAGGAAGACGACCGCTACTACACAGCCATCACTTTTGTGGCTAC		3240
3555	QY	TCCTGAGCAACAAAACAGGATTTGGACATGTTCA	TCAATGCTCTCAAGAAATTCACACT	3614
3241	DB	TCCTGATGAACAAAACAGGATTTGGACATGTTCA	TCAATGCTCTCAAGAAATTCACACT	3300
3615	QY	CAACATCACTCGGCTGCCAGTTTCTCAGCTGGAAACCCAGCTCGGAGAAAGATGCGCTGT		3674
3301	DB	CAACATCACTCGGCAACAGCTTCCAGCCGGAACCCAGCTCGGAGAAAGATGCGCTGT		3360
3675	QY	TGTTTCAAAAACCAACATTAAAGAGGTACAAAGATAGTTTCTCTAATGAGAAATTTGATTT		3734
3361	DB	TGTTTCAAAAACCAACATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTT		3420
3735	QY	TCGCAACACCCAAATATCACTTCTTTGTTTATGTCAAGTAAATTCACCTGCGCCCATCAA		3794
3421	DB	TCGCAACACCATCCAAACATCACTTCTTTGTTTATGTCAAGTAAATTCACCTGCGCCCATCAA		3480
3795	QY	AAATTCAGATTGCGCTTCTCTCAGCAGCAGCAAAATTTATGACCTGGTACAGTTTCTTCGTGAC		3854
3481	DB	AAATTCAGATTGCGCTTCTCCAGCAGCAGCAAAATTCATGACCTGGTACAGTTTCTTCGTGAC		3540
3855	QY	TTTCTTCAGTTGTTTCTCTCTTTGCTCTGTGTGGCTGCTGTGGTTTGGAGAGTCAACA		3914
3541	DB	TTTCTTCAGTTGTTTCTCTTCGCTGCTTCTGTGTGGCTGCTGTGGAGATCAAGCA		3600
3915	QY	AAATTTGTTGGCGCTCCAGACGCTAGAGCAAACTTCTTCGAGAGATGCAACAGATGGCCAG		3974
3601	DB	GAGCTGTTGGCATCCAGCGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGCCAG		3660
3975	QY	CCGTCCCTTTGGCTCTGTAAATGTCGCTTGGAAACAGATGAGGACCTCTCTGATCTTAT		4034
3661	DB	CCGCGCTTTGCTCTGTAAACGTTGCCCTTGGAAACAGATGAAGAACCTCTCTGATCTCAT		3720
4035	QY	TGGGGGAGGTATTAAGACTGTTCCCAAAACCCATTGCACTGGAGCCGCTGTTTGGCAACAA		4094
3721	DB	TGGGGGAGGTATTAAGACCGCTTCTTAAGCCCAATTTGCCCTGGAGCCCTGCTTTGGTAAACAA		3780



1755 QY CTTCTCTTCCAGATTTCATGGCCCTATGACATTTGCTGTGACCGCTGCTCAGTGTCTCCCGAG 1814  
1261 Db CTTCTCTTCCAGATTTCATGGCCCTATGACATTTGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1320  
1815 QY ACTGATCTCCACCATGATGTCAACAGATTGGCCATTGACAGTCTTACACACAGCAC 1874  
1321 Db NNN 1380  
1875 QY CATGTATGTCTGGTGGTGGTTCATAGTCTCTCCCTCAGCGACATCTCGTATTCACCTC 1934  
1381 Db CATGTATGTCTGGTGGTGGTTCATAGTCTCTCCCTCAGCGACATCTCGTATTCACCTC 1440  
1935 QY GGAACAGTGTGATGGCGATCGAGTGAAGCGCTTGTATTAGCAGCAGGACCTGGTATTCG 1994  
1441 Db GGAACAGTGTGATGGCGATCGAGTGAAGCGCTTGTATTAGCAGCAGGACCTGGTATTCG 1500  
1995 QY GTGTGTGTGGAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATCAACA 2054  
1501 Db GTGTGTGTGGAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATCAACA 1560  
2055 QY AGAAGAAAGTTAAATCAGAAATGTTTTTCAAAGAACTCTTGACCATGACAGATGTGA 2114  
1561 Db AGAAGAAAGTTAAATCAGAAATGTTTTTCAAAGAACTCTTGACCATGACAGATGTGA 1620  
2115 QY CCAGCACACAGATTGTACAGCTGTACAGCCACACCAATGACTGCCACTGGTGCAATGA 2174  
1621 Db CCAGCACACAGATTGTACAGCTGTACAGCCACACCAATGACTGCCACTGGTGCAATGA 1680  
2175 QY CCATTGTGTCTCCAGGAAACACAGCTGTCTCAGAGGCCAGATCTCCATTTTATAGTATGA 2234  
1681 Db CCATTGTGTCTCCAGGAAACACAGCTGTCTCAGAGGCCAGATCTCCATTTTATAGTATGA 1740  
2235 QY GAATTTGCCCAAGATTAACCCCATGTACTGTCTAAACAGAGACAGCTGCGAGGAGCTG 2294  
1741 Db GAATTTGCCCAAGATTAACCCCATGTACTGTCTAAACAGAGACAGCTGCGAGGAGCTG 1800  
2295 QY TGCCCTGACACAGAACTGCCAGTGGGAGCCCGGAACTCAGAGTGCATTTGCCCTGCCCGA 2354  
1801 Db TGCNNNGAACACAACTGCCAGTGGGAGCCCGGAACTCAGAGTGCATTTGCCCTGCCCGA 1860  
2355 QY AAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAAAATCTCATGTTTGAATAATCTACTGC 2414  
1861 Db AAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAAAATCTCATGTTTGAATAATCTACTGC 1920  
2415 QY CAAGGAGNATTATGCAANTGTAAATTTCTGTAGGAAACACAAATGCCCTTTTGGGTTTC 2474  
1921 Db CAAGGAGNATTATGCAANTGTAAATTTCTGTAGGAAACACAAATGCCCTTTTGGGTTTC 1980  
2475 QY TCTTACAAACCCAGAGAGGTAGAAATTTGTCTTTAAGCAGCTGCGAATAATGCAAGTCAATC 2534  
1981 Db TCTTACAAACCCAGAGAGGTAGAAATTTGTCTTTAAGCAGCTGCGAATAATGCAAGTCAATC 2040  
2535 QY TCAGAGCATGTCCAAAGTCACTTAACCCCATGGTGGCCCTTCGAGAGATCAATGTGTC 2594  
2041 Db TCNNNNNNNTCCAAAGCTCACCTTAACCCCATGGTGGCCCTTCGAGAGATCAATGTGTC 2100  
2595 QY CTACTGGTGTGGGAAGATATGTCCCATTTTACAAATAGTTTACTACAGTGGATGCCGTC 2654  
2101 Db CTACTGGTGTGGGAAGATATGTCCCATTTTACAAATAGTTTACTACAGTGGATGCCGTC 2160  
2655 QY TGAGCCCAAGTGTGATGTGGAATTTTATGGAATTTTATCAGAAACCGAGTACTCGGAGCTGAA 2714  
2161 Db TGAGCCCAAGTGTGATGTGGAATTTTATGGAATTTTATCAGAAACCGAGTACTCGGAGCTGAA 2220  
2715 QY GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGCCCTGCAAAACACAG 2774  
2221 Db GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGCCCTGCAAAACACAG 2280  
2775 QY TGCTAAGCAGTCCCGACACCAATGTGCCCTTGAAGCAGCATGTGGAGATTGCAACAGCGG 2834  
2281 Db TGCTAAGCAGTCCCGACACCAATGTGCCCTTGAAGCAGCATGTGGAGATTGCAACAGCGG 2340  
2835 QY CAGCTCTGAGTGCATGTGTGAGCAACATGAGCAGTGTGTGACTCCAAATGCCTATGT 2894

2341 Db CAGCTCTGAGTGCATGTGTGTCAGCAACATGAGCAGTGTGTGACTCCAAATGCTCAGT 2400  
2895 QY GGCCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTCCGCCCTCGA 2954  
2401 Db GGCCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTCCGCCCTCGA 2460  
2955 QY AAATTGTTCAAGCTACTGTACTCTGTAGTCAATTTGCTTGAGCAACAGCAGCTGTGGCTG 3014  
2461 Db NNN 2520  
3015 QY TACTGATCCAGCAATACTGGCAAGGGAATGATAGAGGGTTCTATAAAGGACCAGT 3074  
2521 Db NNN 2580  
3075 QY GAAGATGCTTCGCAAGCCCTACAGGAAAATTTCTATCCACAGCCCTGCTCAATCCAG 3134  
2581 Db NNN 2640  
3135 QY CATGTGTCTAGAGGACAGCAGATCAAACTGGTCTTTCAATTCAGTGTCCAGCTTGCATG 3194  
2641 Db NNN 2700  
3195 QY CAAACGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACTGACCAC 3254  
2701 Db CAAACGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACTGACCAC 2760  
3255 QY AGCAAGCAGCTCGGAGACCTGATCTGGCTTTCTACGGTGTATCCCAATGAGGGAA 3314  
2761 Db AGCAAGCAGCTCGGAGACCTGATCTGGCTTTCTATGGTGTATCCCAATGAGGGAA 2820  
3315 QY ATGTGAGCCATCAAGTGCATGGGACAGCGCTCTGTGCAACACAAACAGCGGCAAGTG 3374  
2821 Db NNN 2880  
3375 QY CTTTGTGACCAACAGGGGCTCAAGGGGACAGTGCAGCTATGTGAGTGAAGAAATCG 3434  
2881 Db CTTTGTGACCAACAGGGGCTCAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2940  
3435 QY ATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTTATATCTCTTTATTTGACTATCAGTT 3494  
2941 Db ATACNN 3000  
3495 QY CACCTTTAGTCTATCCAGGAAAGATGATCGCTATTATACAGCTATCAATTTTGTGGCTAC 3554  
3001 Db CACCTTTAGTCTATCCAGGAAAGATGATCGCTATTATACAGCTATTAATTTTGTGGCTAC 3060  
3555 QY TCCTGAGCAACAAACAGGGATTTTGACATGTTTCATCAATGCTCCAGAAATTTCAACCT 3614  
3061 Db TCCTGAGCAACAAACAGGGATTTTGACATGTTTCATCAATGCTCCAGAAATTTCAACCT 3120  
3615 QY CAAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTGT 3674  
3121 Db CAAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTGT 3180  
3675 QY TGTTTCAAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTCAATTT 3734  
3181 Db TGTTTCAAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTCAATTT 3240  
3735 QY TCGCAACCCACCAATATCACTTTCTTTGTTATGTCAGTAATTTTCACTGGGCCATCAA 3794  
3241 Db TCGCAACCCACCAATATCACTTTCTTTGTTATGTCAGTAATTTTCACTGGGCCATCAA 3300  
3795 QY AATTGAGATTGCTTCTCTCAGCAGCAAGCAATTTTATGAGACCTGGTACAGTTCTTCGTGAC 3854  
3301 Db AATTGAGATTGCTTCTCTCAGCAGCAAGCAATTTTATGAGACCTGGTACAGTTCTTCGTGAC 3360  
3855 QY TTTCTTCAGTTGTTTCTCTCTTTTGTCTGCTGGTGTCTGTTGTTGGAAGATCAAAACA 3914  
3361 Db TTTCTTCAGTTGTTTCTCTCTTTTGTCTGCTGGTGTCTGTTGGAAGATCAAAACA 3420  
3915 QY AAGTTGTGGGCTCCAGAGCTAGAGAGCAATTTCTTCGAGAGATGCAACAGATGGCCAG 3974

```

Db      3421 AAGTTGTTGGGCTCCAGACGTAGAGACCACTTCTTCGAGAGATGCAACAGATGGCCAG 3480
QY      3975 CGTCCCTTTGGCTCTGTAATGTGGCTTGGAAACAGATGAGGAGCTCTCTGATCTTAT 4034
Db      3481 CGTCCCTTTGGCTCTGTAATGTGGCTTGGAAACAGATGAGGAGCTCTCTGATCTTAT 3540
QY      4035 TGGGGGGAGTAAAGACTGTTCCCAACCCATTGCACTGGAGCGTGTGTTTGGCAACA 4094
Db      3541 TGGGGGGAGTAAAGACTGTTCCCAACCCATTGCACTGGAGCGTGTGTTTGGCAACA 3600
QY      4095 AGCCGCTGCTCTCTGTTGTTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCTCC 4154
Db      3601 AGCCGCTGCTCTCTGTTGTTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCTCC 3660
QY      4155 TGGGCGAGTCAG 4165
Db      3661 TGGGCGAGTCAG 3671

RESULT 4
LOCUS   BX440935
DEFINITION BX440935 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
          CSODF012YL19 5-PRIME, mRNA sequence.
ACCESSION BX440935
VERSION   BX440935.1 GI:30781889
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1201)
AUTHORS   Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished (2001)
COMMENT   Contact: Genoscope
          Genoscope - Centre National de Sequencage
          BP 191 91006 EVRY cedex - France
          Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
          Library was constructed by Life Technologies, a division of
          Invitrogen. This sequence belongs to sequence cluster 10212.f For
          more information about this cluster, see
          http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CSODF012CF10QPI&cluster=10212.f. Contact :
          Feng Liang Email : fliang@lifetech.com URL :
          http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
          Faraday Avenue Genoscope sequence ID : CSODF012CF10QPI.
FEATURES
         source
         1..1201
         /organism="Homo sapiens"
         /mol_type="mRNA"
         /db_xref="taxon:9606"
         /clone="CSODF012YL19"
         /tissue_type="FETAL BRAIN"
         /dev_stage="fetal"
         /clone_lib="Homo sapiens FETAL BRAIN"
         /notes="Organ: brain; Vector: pCMVSPORT.6; 1st strand cDNA
         was primed with a NotI-oligo(dT) primer. Five prime end
         enriched, double-strand cDNA was digested with Not I and
         cloned into the Not I and EcoRV sites of the pCMVSPORT 6
         vector. Library was not normalized."

Query Match 21.68; Score 925.6; DB 13; Length 1201;
Best Local Similarity 98.4%; Pred. No. 5.5e-166;
Matches 962; Conservative 4; Mismatches 8; Indels 4; Gaps 3;

QY      2337 GTGCATTGCCCTGCCGAAATATCTGTGGCATTTGGTGGCATTTGGTGGAACTCATG 2396
Db      57 GTGCATTGCCCTGCCGAAATATCTGTGGCATTTGGTGGCATTTGGTGGAACTCATG 116
QY      2397 TTTGAAATATCTGCGCAAGGAGATATGACAATGCTAAATGTTCTGTAGGAACCA 2456

```

```

Db      117 TTTGAAATATCTGCGCAAGGAGATATGACAATGCTAAATGTTCTGTAGGAACCA 176
QY      2457 CAATGCCCCCTTTTGGCTTCTCTTACAAACCCAGAGAGGTAGAAATTTGTCTTTAAGCAGCT 2516
Db      177 CAATGCCCCCTTTTGGCTTCTCTTACAAACCCCTGATGAAGGTAGAAATTTGTCTTTAAGCAGCT 236
QY      2517 GCGAATTAATGCACTGATCTCAGAGCATGTCAGAGTCACCTTAAACCCATGCGGTGCGCT 2576
Db      237 GCGAATTAATGCACTGATCTCAGAGCATGTCAGAGTCACCTTAAACCCATGCGGTGCGCT 296
QY      2577 TCGGAAGATCAATGTCCTCTACTGCTGGTGGGAAGATATGTCCTCATTTACAAATAGTTT 2636
Db      297 TCGGAAGATCAATGTCCTCTACTGCTGGTGGGAAGATATGTCCTCATTTACAAATAGTTT 356
QY      2637 ACTACAGTGGATGCCGTCTGAGCCCAAGTATGCTGGATTTCTGTGAAATTTTATCAGAAC 2696
Db      357 ACTACAGTGGATGCCGTCTGAGCCCAAGTATGCTGGATTTCTGTGAAATTTTATCAGAAC 416
QY      2697 CAGTACTCGGGGACTGAAGGCTGCAACCTGCACTCAACCCACTCAATGGTAGTCTGTGA 2756
Db      417 CAGTACTCGGGGACTGAAGGCTGCAACCTGCACTCAACCCACTCAATGGTAGTCTGTGA 476
QY      2757 AAGGCTTGCACCAACCAAGTGTAAAGCAGTCCCGGACACCATGTGCTTGAAGACAGCATG 2816
Db      477 AAGGCTTGCACCAACCAAGTGTAAAGCAGTCCCGGACACCATGTGCTTGAAGACAGCATG 536
QY      2817 TGGAGATTGCACAGGCGGAGCTCTGAGTGCATGTGTTGCGAGCAACATGAGAGCAGTGT 2876
Db      537 TGGAGATTGCACAGGCGGAGCTCTGAGTGCATGTGTTGCGAGCAACATGAGAGCAGTGT 596
QY      2877 GGACTCCCAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACCAT 2936
Db      597 GGACTCCCAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACCAT 656
QY      2937 GAGCACCTGCCCCCTCGAAATTTGTTCAAGGCTACTGTACCTGTAGTCAATGCTTGGAGCA 2996
Db      657 GAGCACCTGCCCCCTCGAAATTTGTTCAAGGCTACTGTACCTGTAGTCAATGCTTGGAGCA 716
QY      2997 ACCAGGCTGTGGCTGTGCTACTGATCCAGCAATCTGCGAAAGGAAATCAATAGAGGG 3056
Db      717 ACCAGGCTGTGGCTGTGCTACTGATCCAGCAATCTGCGAAAGGAAATCAATAGAGGG 776
QY      3057 TTTCTATAAAGACCAAGTGAAGATGCTTCCCAAGCCCTTACAGGAAATTTCTATCCACA 3116
Db      777 TTTCTATAAAGACCAAGTGAAGATGCTTCCCAAGCCCTTACAGGAAATTTCTATCCACA 836
QY      3117 GCGCCCTGCTCAATTCACAGTGTCTAGAGAGCAGAGATACAACTGGCTTTTCATCA 3176
Db      837 GCGCCCTGCTCAATTCACAGTGTCTAGAGAGCAGAGATACAACTGGCTTTTCATCA 896
QY      3177 CTGTCCAGCTTGCCTAATCAACGGCCACAGTAATATGCAATCAATCAGAGCATCTGTGAGAA 3236
Db      897 CTGTCCAGCTTGCCTAATCAACGGCCACAGTAATATGCAATCAATCAGAGCATCTGTGAGAA 956
QY      3237 GTGTGAGAACTGACCAAGCAGGACCTGCGAGACCTGCAATCTGTGGCTTCTACGGTGA 3296
Db      957 GTGTGAGAACTGAGAAC--AMAGGAGACCTGCGAGAACTGCATATCTGCG--TCTACGGKGA 1013
QY      3297 TCCCAACCAATGGAGGAA 3314
Db      1014 T-CCACCAATGGAGGAA 1030

RESULT 5
LOCUS   BU506373
DEFINITION AGSCOURT 10015485 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6494441
          5', mRNA sequence.
ACCESSION BU506373
VERSION   BU506373.1 GI:22812606
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 964)

NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM14050 row: k column: 18  
High quality sequence stop: 592.  
Location/Qualifiers  
1. 964  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:649441"  
/tissue\_type="retina"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 94"  
/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 18.1%; Score 775; DB 13; Length 964;  
Best Local Similarity 88.9%; Pred. No. 3e-137;  
Matches 857; Conservative 0; Mismatches 105; Indels 2; Gaps 2;

QY 1996 TGTGTGGAAACACAGGTCCTCTCAGTGATATCTGTCGGGGCTGGCAACTGTATGAACAA 2055  
DB 1 TGTGTGGGACACACAGTCGTCTCGATGTACCTCTCTGGAGTTGGCAACTGAAGAACAA 60

QY 2056 GAAGAAAGTTAAATCAGATGTTTTCCTCAAGAACTCTTGACCATGACAGATGAC 2115  
DB 61 GCAGAAAGTTAAATCAGATGTTTTCCTCAAGAACTCTTGACCATGACAGATGAC 120

QY 2116 CAGCACAGATTTGTACAGCTGTACAGCCCAACCAATGATGCCATGCTGGTGAATGAC 2175  
DB 121 CAGCACAGATTTGTACAGCTGTACAGCCCAATGATGCCATGCTGGTGAATGAC 180

QY 2176 CATTGTGTCGCCAGGAACACACAGCTGTCTCAGAGCCAGATCTCCATTTTAGTATGAC 2235  
DB 181 CACTGTGTCCTGTGAACACACAGCTGTCTCAGAGCCAGATCTCCATTTTAGTATGAC 240

QY 2236 AATTGCCCAAGGATAACCCATGTACTACTGTATCAAGAGACACAGCTGCAGAGCTGT 2295  
DB 241 AGTTGCCCAAGGATAACCCATGTACTACTGTATCAAGAGACACAGCTGCAGAGCTGT 300

QY 2296 GCCCTGGACCAAGATGCGCAGTGGAGGCCCGGAATCAGAGTGAATGCCCTGCCGAA 2355  
DB 301 GCCCTAGACCAAGATGCGCAGTGGAGGCCCGGAATCAGAGTGAATGCCCTGCCGAA 360

QY 2356 AATATCTGGCATTTGGTGGAACTCATGTTTGAAATTAATTAATTAATTAATTAATTAAT 2415  
DB 361 AATATCTGGCATTTGGTGGAACTCATGTTTGAAATTAATTAATTAATTAATTAATTAAT 420

QY 2416 AAGGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2475  
DB 421 AAGGAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480

QY 2476 CTTACAAACCAAGAGTAGAATTTGCTTACAGCTGCGAATTAATTAATTAATTAATTAAT 2535  
DB 481 CTCACATCCCAAGAGTAGAATTTGCTTACAGCTGCGAATTAATTAATTAATTAATTAAT 540

QY 2536 CAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTTCGGAGATCAATGTGTCC 2595  
DB 541 CAAAGTATGTCCAAGCTCACCTGTGACTCCATGGGTGGTCTTCGGAGATCAATGTGTCT 600

QY 2596 TACTGTCTCTGGGAAGATATGTCTCCCATTTACAATAGTTTACTACAGTGGATCCGTCT 2655  
DB 601 TACTGTCTCTGGGAAGATATGTCTCCCATTTACAATAGTTTACTACAGTGGATCCGTCT 660

QY 2656 GAGCCCACTGTATGCTGGGATTTCTGGGAATTTATCAGAACCCAGTACTC-GGGAGCTGAA 2714  
DB 661 GAGCCCACTGTATGCTGGGATTTCTGGGATTTATCAGAACCCAGTACTC-GGGAGCTGAA 720

QY 2715 GGCTCAACCTGTATGCTCAACCCACTCAATGTAGTGTCTGTGAAGGCTGCAAAACACAG 2774  
DB 721 GGCTCAACCTGTATGCTCAACCCACTCAATGTAGTGTCTGTGAAGGCTGCAAAACACAG 780

QY 2775 TGCTAAGCAGTGCCTGGACACCATGTGCTTGGAGACAG-CATGTGGAGATTGCACACGC 2833  
DB 781 TGCTAAGCAGTGCCTGGACACCATGTGCTTGGAGACAG-CATGTGGAGATTGCACACGC 840

QY 2834 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAGACAGTGTGTGACTCCATGCTATG 2893  
DB 841 CCAGCTCGGAGTGCATGTGGTGCAGCAACATGAGACAGTGTGTGACTCCATGCTATG 900

QY 2894 TGGCT 2953  
DB 901 TGGCT 960

QY 2954 AAAA 2957  
DB 961 AAAA 964

RESULT 6  
BG678679 933 bp mRNA linear EST 01-MAY-2001  
LOCUS 602624477F1 NCI\_CGAP\_Skn4 Homo sapiens cDNA clone IMAGE:4749427 5',  
DEFINITION mRNA sequence.  
ACCESSION BG678679.1 GI:13910076  
VERSION BG678679.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 933)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgabs@mail.nih.gov](mailto:cgabs@mail.nih.gov)  
Tissue Procurement: James Cleaver, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLAM10602 row: j column: 20  
High quality sequence stop: 762.  
Location/Qualifiers  
1. 933  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4749427"  
/tissue\_type="squamous cell carcinoma"  
/lab\_host="DH10B (TI phage-resistant)"  
/clone\_lib="NCI CGAP Skn4"  
/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: oligo dr. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI\_CGAP Library."

FEATURES  
source



## ORIGIN

Query Match 17.7%; Score 758.8; DB 12; Length 933;  
 Best Local Similarity 98.4%; Pred. No. 3.6e-134; Indels 1; Gaps 1;  
 Matches 777; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

2997 ACAGGCTGTGGCTGTGTACTGATCCAGCAATATCTGCAAGGGAATGCATAGAGGG 3056  
 Db 2 ACAGGCTGTGGCTGTGTACTGATCCAGCAATATCTGCAAGGGAATGCATAGAGGG 61

3057 TTCCTATAAGGACGAGTGAAGTGCCTTCGCAAGCCCTACAGAAATTTCTATCCACA 3116  
 Db 62 TTCCTATAAGGACGAGTGAAGTGCCTTCGCAAGCCCTACAGAAATTTCTATCCACA 121

3117 GCCCTGTCTCAATTCAGCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTTCATCA 3176  
 Db 122 GCCCTGTCTCAATTCAGCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTTCATCA 181

3177 CTGTCCAGTTCCCAATGCAAGCGCCACAGTAATATGATCAATCAGACATCTGTGAAA 3236  
 Db 182 CTGTCCAGTTCCCAATGCAAGCGCCACAGTAATATGATCAATCAGACATCTGTGAAA 241

3237 GTGTGAGAACCTTGACACAGGCAAGCACTGGGAGACCTGCAATATCTGGCTTCTACGGTGA 3296  
 Db 242 GTGTGAGAACCTTGACACAGGCAAGCACTGGGAGACCTGCAATATCTGGCTTCTACGGTGA 301

3297 TCCCAACCAATGAGGGAATGTGAGCCATGCAAGTGCATGCGCAAGCGCTCTGTGCAA 3356  
 Db 302 TCCCAACCAATGAGGGAATGTGAGCCATGCAAGTGCATGCGCAAGCGCTCTGTGCAA 361

3357 CACCAACAGGCAAGTGTCTGTGACCAACAGGCGGTCAAGGGGACGATCGCAAGCT 3416  
 Db 362 CACCAACAGGCAAGTGTCTGTGACCAACAGGCGGTCAAGGGGACGATCGCAAGCT 421

3417 ATGTGAGTGAAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATATCTCT 3476  
 Db 422 ATGTGAGTGAAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATATCTCT 481

3477 TCTTATTCAGTATCAGTTCAGTTTACTTCTAGTCTATCCAGGAAGATGCTATTACAGC 3536  
 Db 482 TCTTATTCAGTATCAGTTCAGTTTACTTCTAGTCTATCCAGGAAGATGCTATTACAGC 541

3537 TATCAATTTTGTGGCTACTCTGACGAAACCAACAGGATTTGACATGTTTCATCAATGC 3596  
 Db 542 TATCAATTTTGTGGCTACTCTGACGAAACCAACAGGATTTGACATGTTTCATCAATGC 601

3597 CTCGAAGATTTCAACCTCAACATCAGCTGGGCTGCGAGTTTCTCAGCTGGAACCCAGGC 3656  
 Db 602 CTCGAAGATTTCAACCTCAACATCAGCTGGGCTGCGAGTTTCTCAGCTGGAACCCAGGC 661

3657 TGGAGAGAGATGCTGTGTTTTCAGAAACCAACATTAAGAGTACAAAGATAGTTCTC 3716  
 Db 662 TGGAGAGAGATGCTGTGTTTTCAGAAACCAACATTAAGAGTACAAAGATAGTTCTC 721

3717 TAATGAGAAGTTGATTTTTCGCAACCAACCAATATCATCTTCTTTGTTTATGTCAGTAA 3776  
 Db 722 TAATGAGAAGTTGATTTTTCGCAACCAACCAATATCATCTTCTTTGTTTATGTCAGTAA 780

3777 TTTTCACTGG 3786  
 Db 781 TTGTCACTTG 790

RESULT 7  
 CB520374  
 LOCUS CB520374  
 DEFINITION UT-M-G10-cei-1-18-0-UI.r1 NIH-BMAP\_G10 Mus musculus cDNA clone  
 IMAGE:6840067 5', mRNA sequence.  
 ACCESSION CB520374  
 VERSION CB520374.1 GI:29353729  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 840)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-r@mail.nih.gov  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
 http://genome.uiowa.edu/distribution/mousefl.html  
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

Seq primer: pYX-5.

Location/Qualifiers  
 1. .840  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:6840067"  
 /issue\_type="whole brain"  
 /dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH BMAP G10"  
 /note="Organ: Brain; Vector: pYX-Asc; Site\_1: EcoR I; Site\_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is ACCGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

## ORIGIN

Query Match 16.3%; Score 701.2; DB 14; Length 840;  
 Best Local Similarity 90.2%; Pred. No. 3.5e-123;  
 Matches 759; Conservative 0; Mismatches 81; Indels 1; Gaps 1;

2151 CAATGACTGCCACTGGTGCATGACCATGTGTCCCAAGAACACACACTGCTCAGAGG 2210  
 Db 1 CAATGACTGCCACTGGTGCATGACCATGTGTCCCTGTGAACACACACTGCAAGAGG 60

2211 CCAGATCTCCATTTTAGGTATGAAATGCCCCAAGGATAACCCATGTACTACTGTAA 2270  
 Db 61 CCAGATCTCCATTTAGGTATGAGAGTTGCCCAAGGATAACCCATGTACTACTGCAA 120

2271 CAGAGACAGCTGTCAGAGCTGTGCCCTGGACAGAACTGCCAGTGGGAGCCCCGAA 2330  
 Db 121 TAAGAAAACAGCTGTCAGAGCTGTGCCCTAGACAGAACTGCCAGTGGGAGCCCCGAA 180

2331 TCAGAGTGCATTCGCCCTGCCGAAATATCTGTGGCAATTCGCTGGTGGTGGTGGAA 2390  
 Db 181 TCAGAGTGCATTCGCCCTGCCGAAATATCTGTGGCAATTCGCTGGTGGTGGTGGAA 240

2391 CTCATGTTTGAATTTACTACTGCAAGGAAATATGCAATGCTAAATTTGTTCTAG 2450  
 Db 241 CTCGTGTCTGAAATCACTACTGTAAGGAGAAATATGCAATGCTAAATTTGTTCTAG 300

2451 GAACCAATGCTTTTGGCTTCTCTTACACCCCAAGAGTGAATTTGTCCTTAA 2510  
 Db 301 GAACCAATGCTTTTGGCTTCTCTTACACCCCAAGAGTGAATTTGTCCTTAA 360

QY 2511 GCAGCTGGCAATAAGTCAGTCACTCAGAGCATGTCAGAGCTCACTTAACCCCATGGGT 2570  
 Db 361 GCAGCTTCGATTAATGCAATCAATCAATCAATGATGTCAGAGCTCACTGACCTCCATGGGT 420  
 QY 2571 CGGCCCTCGAAGATCAATATGTCTTACTTGGTCTGGAGAGATATGTCCCATTTACAAA 2630  
 Db 421 TGGTCTTCGGAAGATCAATATGTCTTACTTGGTCTGGAGAGATATGTCTCCATTCACAAA 480  
 QY 2631 TAGTTTACTACAGTGGATCGCTGAGCCAGTATGCTGGATCTGTGGATTTTATC 2690  
 Db 481 TAGTTTGTCTGCAAGTGGATGCAATCTGAGCCAGTATGCTGGATTTTATC 540  
 QY 2691 AGAACCCAGTACTCGGGGAGTGAAGCTGCAACCTGATCAACCCCACTCAATGGTAGTGT 2750  
 Db 541 AGAGCCCTAGTACTCGGGGATTAAGGCTGAACCTGATCAACCCCTCTCAATGGCAGGT 600  
 QY 2751 CTGTGAAAGCCCTGCAACACAGTCTTAAGCAAGTCCCGACACCATGTGCTTGGAGAC 2810  
 Db 601 CTGTGAAAGCCCTGCAACACAGTCTTAAGCAAGTCCCGACACCATGTGCTTGGAGAC 660  
 QY 2811 AGCATGTGAGATTCACAGCGGAGCTCTGAGTGCATGTGTGAGCAACATGAAGCA 2870  
 Db 661 AGCGTGTGGAGTGCATAGCAGAGCTCGGAGTGCATGTGTGAGCAACATGAAGCA 720  
 QY 2871 GTGTGTGAGTCCCAATGCTTGTGAGTGTGCTTGTGAGTGTGAGTGTGAGTGTGAGTGT 2930  
 Db 721 GTGTGTGAGTCCCAATGCTTGTGAGTGTGCTTGTGAGTGTGAGTGTGAGTGTGAGTGT 780  
 QY 2931 TAGATGAGCACTGCCCCCTGAAATTTGTTGAGGCTACTGACCTGTGAGTGTGAGTGTGAGTGT 2990  
 Db 781 TAGATGAGCACTGCCCCCTGAAATTTGTTGAGGCTACTGACCTGTGAGTGTGAGTGTGAGTGT 839  
 QY 2991 G 2991  
 Db 840 G 840

RESULT 8  
 BM783739  
 LOCUS K-EST0061730 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-H05 5',  
 DEFINITION mRNA sequence.

ACCESSION BM783739  
 VERSION BM783739.1 GI:19131971  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 881)  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 34 row: H column: 05  
 High quality sequence stop: 681.  
 Location/Qualifiers

FEATURES  
 source  
 1..681  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S5SNU484-34-H05"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Epithelial"

/cell\_line="SNU-484"  
 /lab\_host="Top10F"  
 /clone\_lib="S5SNU484"  
 /note="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;  
 Site 2: NotI. The poly (A) + RNA was decapped with tobacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 RNA ligase. The  
 first strand cDNA was synthesized from oligo dt-tailed  
 mRNA by priming with dt-tailed vector. The cDNA vector  
 was adjusted to have about 60nt. The cDNA vector  
 was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a DNA strand by Okayama-Berg method.  
 The obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F by electroporation  
 method."

ORIGIN

Query Match 15.9%; Score 681; DB 12; Length 681;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-119; Indels 0; Gaps 0;  
 Matches 681; Conservative 0; Mismatches 0;  
 QY 2512 CAGCTTGCAGATTAATGCACTCACTCAGAGCATGTCAGAGCTCACTTAACCCCATGGGT 2571  
 Db 1 CAGCTTGCAGATTAATGCACTCACTCAGAGCATGTCAGAGCTCACTTAACCCCATGGGT 60  
 QY 2572 GGCCTTCGGAAGATCAATATGTCTTACTTGGTCTGGAGAGATATGTCCCATTTACAAAT 2631  
 Db 61 GGCCTTCGGAAGATCAATATGTCTTACTTGGTCTGGAGAGATATGTCCCATTTACAAAT 120  
 QY 2632 AGTTTACTACAGTGGATGCCCTCTGAGCCAGTGCATGTGAGTCTGTGGAATTTATCA 2691  
 Db 121 AGTTTACTACAGTGGATGCCCTCTGAGCCAGTGCATGTGAGTCTGTGGAATTTATCA 180  
 QY 2692 GAACCCAGTACTCGGGGACTGAAGGCTCAACCTGTGATCAACCCCACTCAATGGTAGTGT 2751  
 Db 181 GAACCCAGTACTCGGGGACTGAAGGCTCAACCTGTGATCAACCCCACTCAATGGTAGTGT 240  
 QY 2752 TGTGAAAGCCCTGCAACACAGTCTTAAGCAGTCCCGACACCATGTGCTTGGAGCA 2811  
 Db 241 TGTGAAAGCCCTGCAACACAGTCTTAAGCAGTCCCGACACCATGTGCTTGGAGCA 300  
 QY 2812 GCATGTGGAGATTGCAACAGCGGAGCTCTGAGTGCATGTGTGAGCAACATGAAGCAG 2871  
 Db 301 GCATGTGGAGATTGCAACAGCGGAGCTCTGAGTGCATGTGTGAGCAACATGAAGCAG 360  
 QY 2872 TGTGTGAGTCAATGCCCTATGTGCGCTTCCCTTTGGCCAGTGTGATGAATGGTAT 2931  
 Db 361 TGTGTGAGTCAATGCCCTATGTGCGCTTCCCTTTGGCCAGTGTGATGAATGGTAT 420  
 QY 2932 ACGATGAGCACCTGCCCCCTGAAAATTTGTTGAGGCTACTGTACCTGTAGTCAATGCTTG 2991  
 Db 421 ACGATGAGCACCTGCCCCCTGAAAATTTGTTGAGGCTACTGTACCTGTAGTCAATGCTTG 480  
 QY 2992 GAGCAACAGGCTGTGGTGTGATCTGATCCAGCATCTGCAAGAGGGAATGCATA 3051  
 Db 481 GAGCAACAGGCTGTGGTGTGATCTGATCCAGCATCTGCAAGAGGGAATGCATA 540  
 QY 3052 GAGGGTTCCTATTAAGGACCACTGAAGTGCCTTCGCAAGCCCTACACGAAATTTCTAT 3111  
 Db 541 GAGGGTTCCTATTAAGGACCACTGAAGTGCCTTCGCAAGCCCTACACGAAATTTCTAT 600  
 QY 3112 CCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGCTTTTC 3171  
 Db 601 CCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGCTTTTC 660  
 QY 3172 ATTCACTGTCAGCTTGCCAA 3192  
 Db 661 ATTCACTGTCAGCTTGCCAA 681

RESULT 9  
 BM783788

```

LOCUS      BM783788      683 bp      mRNA      linear      EST 05-MAR-2002
DEFINITION K-EST00611794 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-E09 5',
            mRNA sequence.
ACCESSION  BM783788
VERSION     BM783788.1  GI:19132020
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 683)
AUTHORS     Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
TITLE       21C Frontier Korean EST Project 2001
JOURNAL     Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 36 row: E column: 09
            High quality sequence stop: 683.
            Location/Qualifiers
                1..683
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="S5SNU484-36-E09"
                /sex="M"
                /tissue_type="Stomach"
                /cell_type="Epithelial"
                /lab_host="Top10P"
                /clone_lib="S5SNU484"
                /note="Organ: Stomach; Vector: pTZ18RPI; Site:1: EcoRI;
                Site:2: NotI; The poly (A)+ RNA was deapped with tabacco
                acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                including EcoRI site by treatment of T4 RNA ligase. The
                first strand cDNA was synthesized from oligo dT-selected
                mRNA by priming with dT-tailed vector. The dT-tailed
                vector was adjusted to have about 60nt. The cDNA vector
                was circularized with E. coli DNA ligase after digestion
                of EcoRI which site is also included in vector. An RNA
                strand converted to a DNA strand by Okayama-Berg method.
                The obtained cDNA vectors were used for transformation of
                competent cells E. coli Top10P' by electroporation
                method."

ORIGIN
Query Match 15.9%; Score 680.4; DB 12; Length 683;
Best Local Similarity 99.7%; Pred.No. 3.1e-119;
Matches 681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2512 CAGCTGGGAATATGCAAGTCTATCTCAGAGCATGTCGAAGTCCACCTTAACCCCATGGTGC 2571
Db 1 CAGCTGGGAATATGCAAGTCTATCTCAGAGCATGTCGAAGTCCACCTTAACCCCATGGTGC 60

QY 2572 GGCCTTCGGAAGATCAATGTCTCTACTGTGTGGGAAGATATGTCCTCAATTAACAAT 2631
Db 61 GGCCTTCGGAAGATCAATGTCTCTACTGTGTGGGAAGATATGTCCTCAATTAACAAT 120

QY 2632 AGTTTACTACAGTGGATCGCTCTGAGCCAGATGTCGGATTCGTGGAAATTTATCA 2691
Db 121 AGTTTACTACAGTGGATCGCTCTGAGCCAGATGTCGGATTCGTGGAAATTTATCA 180

QY 2692 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTAGTGTGC 2751
Db 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTAGTGTGC 240

QY 2752 TGTGAAGGCTCGAACAACACAGTGTAGACAGTCCGGACACCATGTCCTTGAGGACA 2811

```

```

Db 241 TGTGAAGGCTCGAACAACACAGTGTAGACAGTCCGGACACCATGTCCTTGAGGACA 300
QY 2812 GCATGTGGAGATTGCAACGCGCAGCTCTGAGTGCATGTGTGCGAACCAATGAAGCAG 2871
Db 301 GCATGTGGAGATTGCAACGCGCAGCTCTGAGTGCATGTGTGCGAACCAATGAAGCAG 360
QY 2872 TGTGTGGAGTCCAAATGCTATGTCGCTCTCTCCCTTTTGGCCAGTATGAAGTGTAT 2931
Db 361 TGTGTGGAGTCCAAATGCTATGTCGCTCTCTCCCTTTTGGCCAGTATGAAGTGTAT 420
QY 2932 ACATGAGCAGCTGCCCCCTGAAATTTGTTCCAGGCTACTGTACCTGTAGTCAATGCTTG 2991
Db 421 ACATGAGCAGCTGCCCCCTGAAATTTGTTCCAGGCTACTGTACCTGTAGTCAATGCTTG 480
QY 2992 GAGCAACGAGCTGTGGCTGTGTACTGATCCAGCAATATCTGGCAAGGGAATGCATA 3051
Db 481 GAGCAACGAGCTGTGGCTGTGTACTGATCCAGCAATATCTGGCAAGGGAATGCATA 540
QY 3052 GAGGGTTCCTATATAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTAT 3111
Db 541 GAGGGTTCCTATATAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTAT 600
QY 3112 CCACAGCCCTCTCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACCACTGTGTTTC 3171
Db 601 CCACAGCCCTCTCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACCACTGTGTTTC 660
QY 3172 ATTCACTGTCCAGCTTGCCAATG 3194
Db 661 ATTCACTGTCCAGCTTGCCAATG 683

RESULT 10
LOCUS     BM783258
DEFINITION K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5',
            mRNA sequence.
ACCESSION  BM783258
VERSION     BM783258.1  GI:19131490
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
            21C Frontier Korean EST Project 2001
            Unpublished (2002)
            Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 31 row: H column: 08
            High quality sequence stop: 680.
            Location/Qualifiers
                1..680
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="S5SNU484-31-H08"
                /sex="M"
                /tissue_type="Stomach"
                /cell_type="Epithelial"
                /cell_line="SNU-484"
                /lab_host="Top10P"
                /clone_lib="S5SNU484"
                /note="Organ: Stomach; Vector: pTZ18RPI; Site:1: EcoRI;
                Site:2: NotI; The poly (A)+ RNA was deapped with tabacco
                acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
                including EcoRI site by treatment of T4 RNA ligase. The
                first strand cDNA was synthesized from oligo dT-selected
                mRNA by priming with dT-tailed vector. The dT-tailed
                vector was adjusted to have about 60nt. The cDNA vector
                was circularized with E. coli DNA ligase after digestion
                of EcoRI which site is also included in vector. An RNA
                strand converted to a DNA strand by Okayama-Berg method.
                The obtained cDNA vectors were used for transformation of
                competent cells E. coli Top10P' by electroporation
                method."

FEATURES
source

```

acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-tailed mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transfection of competent cells E. coli Top10F' by electroporation method."

ORIGIN	Query Match	15.9%;	Score 680;	DB 12;	Length 680;
	Best Local Similarity	100.0%;	Pred. No. 3.6e-119;		
	Matches 680;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	2512	CAGCTGCGAATAATGCAAGTATCTCAGAGCATGCCAAGCTCACCTTAACCCCATGGGTC	2571		
Db	1	CAGCTGCGAATAATGCAAGTATCTCAGAGCATGCCAAGCTCACCTTAACCCCATGGGTC	60		
Qy	2572	GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGCCCATTTACAAAT	2631		
Db	61	GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGCCCATTTACAAAT	120		
Qy	2632	AGTTTACTACAGTGGATGCGCTGAGCCAGTCATGCTGGATTCGTGGAAATTTATCA	2691		
Db	121	AGTTTACTACAGTGGATGCGCTGAGCCAGTCATGCTGGATTCGTGGAAATTTATCA	180		
Qy	2692	GAACCCAGTACTCGGGAAGTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	2751		
Db	181	GAACCCAGTACTCGGGAAGTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	240		
Qy	2752	TGTGAAGGCTGCAAAACCAAGTCAAGTGGCGGACACCATGCTGCTTGAGGACA	2811		
Db	241	TGTGAAGGCTGCAAAACCAAGTCAAGTGGCGGACACCATGCTGCTTGAGGACA	300		
Qy	2812	GCATGTGGAGATTGCACCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG	2871		
Db	301	GCATGTGGAGATTGCACCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG	360		
Qy	2872	TGTGTGAGTCAATGCCTATGTGGCTTCCTTCCCTTTGGCCAGTGTATGGATGTAT	2931		
Db	361	TGTGTGAGTCAATGCCTATGTGGCTTCCTTCCCTTTGGCCAGTGTATGGATGTAT	420		
Qy	2932	ACGATGACACTGCCCCCTGAAATGTTCAGGCTACTGTACTGTAGTCATTGCTTG	2991		
Db	421	ACGATGACACTGCCCCCTGAAATGTTCAGGCTACTGTACTGTAGTCATTGCTTG	480		
Qy	2992	GAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATATCGCAAGGGAATGCATA	3051		
Db	481	GAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATATCGCAAGGGAATGCATA	540		
Qy	3052	GAGGTTTCTTAAAGGACCAAGTGAAGTGCCTTGGCAAGCCCTACAGGAAATTTCTAT	3111		
Db	541	GAGGTTTCTTAAAGGACCAAGTGAAGTGCCTTGGCAAGCCCTACAGGAAATTTCTAT	600		
Qy	3112	CCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATCAACTGCTTTTC	3171		
Db	601	CCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATCAACTGCTTTTC	660		
Qy	3172	ATTCACTGTCCAGTTGCCA	3191		
Db	661	ATTCACTGTCCAGTTGCCA	680		

RESULT 11  
BM783773  
LOCUS  
DEFINITION  
K-EST0061775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5',  
rRNA sequence.  
ACCESSION  
BM783773  
VERSION  
BM783773.1 GI:19132005

KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 680)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 36 row: D column: 02  
High quality sequence stop: 680.

FEATURES  
source

1..680  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S5SNU484-36-D02"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/cell\_line="SNU-484"  
/lab\_host="Top10F"  
/clone\_lib="S5SNU484"  
/note="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;  
Site 2: NotI; The poly (A) + RNA was decapped with tobacco  
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
including EcoRI site by treatment of T4 RNA ligase. The  
first strand cDNA was synthesized from oligo dt-tailed  
mRNA by priming with dt-tailed vector. The dt-tailed  
vector was adjusted to have about 60nt. The cDNA vector  
was circularized with E. coli DNA ligase after digestion  
of EcoRI which site is also included in vector. An RNA  
strand converted to a DNA strand by Okayama-Berg method.  
The obtained cDNA vectors were used for transfection of  
competent cells E. coli Top10F' by electroporation  
method."

ORIGIN

Query Match	15.9%;	Score 680;	DB 12;	Length 680;
Best Local Similarity	100.0%;	Pred. No. 3.6e-119;		
Matches 680;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	2512	CAGCTGCGAATAATGCAAGTATCTCAGAGCATGCCAAGCTCACCTTAACCCCATGGGTC	2571	
Db	1	CAGCTGCGAATAATGCAAGTATCTCAGAGCATGCCAAGCTCACCTTAACCCCATGGGTC	60	
Qy	2572	GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGCCCATTTACAAAT	2631	
Db	61	GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGCCCATTTACAAAT	120	
Qy	2632	AGTTTACTACAGTGGATGCGCTGAGCCAGTCATGCTGGATTCGTGGAAATTTATCA	2691	
Db	121	AGTTTACTACAGTGGATGCGCTGAGCCAGTCATGCTGGATTCGTGGAAATTTATCA	180	
Qy	2692	GAACCCAGTACTCGGGAAGTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	2751	
Db	181	GAACCCAGTACTCGGGAAGTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	240	
Qy	2752	TGTGAAGGCTGCAAAACCAAGTCAAGTGGCGGACACCATGCTGCTTGAGGACA	2811	
Db	241	TGTGAAGGCTGCAAAACCAAGTCAAGTGGCGGACACCATGCTGCTTGAGGACA	300	
Qy	2812	GCATGTGGAGATTGCACCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG	2871	

Db 301 GCATGTGGAGATTGACACGCGGCGAGCTGTGAGTCATGTGGTGACAGCAATGAAGCAG 360

QY 2872 TGTGTGGAGTCCAATGCTTATGTGGCCCTCCCTCCCTTTTGGCCAGTGTATGAATGGTAT 2931

Db 361 TGTGTGGAGTCCAATGCTTATGTGGCCCTCCCTCCCTTTTGGCCAGTGTATGAATGGTAT 420

QY 2932 ACATGAGCAGCTGCGCCCTGAAATTTGTCAGGCTACTGTACTGTAGTCATTTGCTTG 2991

Db 421 ACATGAGCAGCTGCGCCCTGAAATTTGTCAGGCTACTGTACTGTAGTCATTTGCTTG 480

QY 2992 GAGCAACAGGCTGGCTGGTGTACTGATCCAGCAATATGCGCAAGGCAAAATGCATA 3051

Db 481 GAGCAACAGGCTGGCTGGTGTACTGATCCAGCAATATGCGCAAGGCAAAATGCATA 540

QY 3052 GAGGTTCTTATAAAGGACCAAGTGAAGTGGCTTGCAGCCCTTACAGGAAATTTCTAT 3111

Db 541 GAGGTTCTTATAAAGGACCAAGTGAAGTGGCTTGCAGCCCTTACAGGAAATTTCTAT 600

QY 3112 CCACAGCCCTGCTCAATTCAGCAGTGTCTAGAGGACAGCAGATCAACTGCTCTTC 3171

Db 601 CCACAGCCCTGCTCAATTCAGCAGTGTCTAGAGGACAGCAGATCAACTGCTCTTC 660

QY 3172 ATTCACTGTCCAGCTTGCCA 3191

Db 661 ATTCACTGTCCAGCTTGCCA 680

RESULT 12

BM783718

LOCUS

DEFINITION K-EST0061702 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-F01 5', mRNA linear EST 05-MAR-2002

ACCESSION BM783718

VERSION BM783718.1 GI:19131950

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 34 row: F column: 01

High quality sequence stop: 677.

Location/Qualifiers

1. .677

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S5SNU484-34-F01"

/sex="M"

/tissue\_type="Stomach"

/cell\_type="Epithelial"

/cell\_line="SNU-484"

/lab\_host="Top10P"

/clone\_lib="S5SNU484"

/note="Organ: Stomach; Vector: pTZ189P1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector

was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10P' by electroporation method."

ORIGIN

Query Match 15.8%; Score 677; DB 12; Length 677;

Best Local Similarity 100.0%; Pred. No. 1.4e-118;

Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2512 CAGCTCGAATATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC 2571

Db 1 CAGCTCGAATATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC 60

QY 2572 GGCTTCGGAAGATCAATGTGCTTACTGCTGGGAAGATATGCTCCCATTTACAAT 2631

Db 61 GGCTTCGGAAGATCAATGTGCTTACTGCTGGGAAGATATGCTCCCATTTACAAT 120

QY 2632 AGTTTACTACAGTGGATGCCGTCTGAGCCCACTGATGCTGGAAATTTATCA 2691

Db 121 AGTTTACTACAGTGGATGCCGTCTGAGCCCACTGATGCTGGAAATTTATCA 180

QY 2692 GAACCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCCACTCAATGGTAGTTC 2751

Db 181 GAACCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCCACTCAATGGTAGTTC 240

QY 2752 TGTGAAAGCCCTGCAAAACACACAGTGTCTAAGCAGTCCCGACACCATGTGCTTGAAGACA 2811

Db 241 TGTGAAAGCCCTGCAAAACACACAGTGTCTAAGCAGTCCCGACACCATGTGCTTGAAGACA 300

QY 2812 GCATGTGGAGATTGACACGCGGCGAGCTGTGAGTCATGTGGTGACAGCAATGAAGCAG 2871

Db 301 GCATGTGGAGATTGACACGCGGCGAGCTGTGAGTCATGTGGTGACAGCAATGAAGCAG 360

QY 2872 TGTGTGGACTCCAATGCTTATGTGGCTTCTCCCTTTTGGCCAGTGTATGGAATGGTAT 2931

Db 361 TGTGTGGACTCCAATGCTTATGTGGCTTCTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

QY 2932 ACATGAGCAGCTGCGCCCTGAAATTTGTCAGGCTACTGTACTGTAGTCATTTGCTTG 2991

Db 421 ACATGAGCAGCTGCGCCCTGAAATTTGTCAGGCTACTGTACTGTAGTCATTTGCTTG 480

QY 2992 GAGCAACAGGCTGGCTGGTGTACTGATCCAGCAATATGCGCAAGGCAAAATGCATA 3051

Db 481 GAGCAACAGGCTGGCTGGTGTACTGATCCAGCAATATGCGCAAGGCAAAATGCATA 540

QY 3052 GAGGTTCTTATAAAGGACCAAGTGAAGTGGCTTGCAGCCCTTACAGGAAATTTCTAT 3111

Db 541 GAGGTTCTTATAAAGGACCAAGTGAAGTGGCTTGCAGCCCTTACAGGAAATTTCTAT 600

QY 3112 CCACAGCCCTGCTCAATTCAGCAGTGTCTAGAGGACAGCAGATCAACTGCTCTTC 3171

Db 601 CCACAGCCCTGCTCAATTCAGCAGTGTCTAGAGGACAGCAGATCAACTGCTCTTC 660

QY 3172 ATTCACTGTCCAGCTTG 3188

Db 661 ATTCACTGTCCAGCTTG 677

RESULT 13

BM783821

LOCUS

DEFINITION K-EST0061830 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-H09 5', mRNA linear EST 05-MAR-2002

ACCESSION BM783821

VERSION BM783821.1 GI:19132053

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE 1 (bases 1 to 677)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 36, row: H column: 09
High quality sequence stop: 677.
Location/Qualifiers
1..677
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SSNU484-36-H09"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="SSNU484"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-tailed
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F by electroporation
method."

ORIGIN
Query Match 15.7%; Score 674.4; DB 12; Length 677;
Best Local Similarity 99.7%; Pred. No. 4.3e-118;
Matches 675; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2512 CAGCTGCGAATATGACGTCTCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGGTC 2571
DB |||||
QY 2572 GGCCTTCGGAAGATCAATGTGCTCTACTGGTGCTGGGAAGATATGTCCTCCATTTTACAAT 2631
DB 61 GGCCTTCGGAAGATCAATGTGCTCTACTGGTGCTGGGAAGATATGTCCTCCATTTTACAAT 120
QY 2632 AGTTACTACAGTGATGCCGTCGTAGCCAGTGTGCTGATCTGTGGATTTTATCA 2691
DB 121 AGTTACTACAGTGATGCCGTCGTAGCCAGTGTGCTGATCTGTGGATTTTATCA 180
QY 2692 GAACCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCTGATGCTGAGCAACCCACTCAATGGTAGTGTC 2751
DB 181 GAACCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCTGATGCTGAGCAACCCACTCAATGGTAGTGTC 240
QY 2752 TGTGAAGCCCTGCACAAACACAGTGTGAAGAGTGCAGGACACCATGTCCTTGAAGACA 2811
DB 241 TGTGAAGCCCTGCACAAACACAGTGTGAAGAGTGCAGGACACCATGTCCTTGAAGACA 300
QY 2812 GCATGTGGAGATGACACGAGCGAGCTCTGATGTCATGTGTGTCAGCAACATGAAGCAG 2871
DB 301 GCATGTGGAGATGACACGAGCGAGCTCTGATGTCATGTGTGTCAGCAACATGAAGCAG 360
QY 2872 TGTGTGGAATCCAAATGCCTATGTGGCCCTCTTCCCTTTGGCCAGTATGGAATGGTAT 2931
DB 361 TGTGTGGAATCCAAATGCCTATGTGGCCCTCTTCCCTTTGGCCAGTATGGAATGGTAT 420

```

```

QY 2932 ACGATCAGACCTGCCCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCAATGCTTG 2991
DB |||||
QY 421 ACGATGAGACCTGCCCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCAATGCTTG 480
DB |||||
QY 2992 GAGCAACCAAGGCTGTGGCTGTACTGATCCCAATATCTGCGAAGCCCTACAGGAATTTCTAT 3051
DB 481 GAGCAACCAAGGCTGTGGCTGTACTGATCCCAATATCTGCGAAGCCCTACAGGAATTTCTAT 540
QY 3052 GAGGGTTCCTATAAGGACCACTGAAGTGTCTTCCGAGCCCTACAGGAATTTCTAT 3111
DB 541 GAGGGTTCCTATAAGGACCACTGAAGTGTCTTCCGAGCCCTACAGGAATTTCTAT 600
QY 3112 CCACAGCCCTCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 3171
DB 601 CCACAGCCCTCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660
QY 3172 ATTCACTGTCAGCTTG 3188
DB 661 ATTCACTGTCAGCTTG 677

RESULT 14
BM783685
LOCUS K-EST0061659 SSNU484 Homo sapiens cDNA clone SSSNU484-34-B06 5',
DEFINITION mRNA sequence.
ACCESSION BM783685
VERSION 1 GI:19131917
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 673)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34, row: B column: 06
High quality sequence stop: 673.
Location/Qualifiers
1..673
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SSNU484-34-B06"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="SSNU484"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transfection of
competent cells E. coli Top10F by electroporation
method."

```

```
ORIGIN
Query Match 15.7%; Score 673; DB 12; Length 673;
Best Local Similarity 100.0%; Pred. No. 7.9e-118;
Matches 673; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2512 CAGCTGCGAATATGAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAAACCCCATGGTGC 2571
Db 1 CAGCTGCGAATATGAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAAACCCCATGGTGC 60

QY 2572 GGCCTTCGGAAGATCAATGTGCTTACTGCTGGAGATATGTCCTCCCATTTACAAT 2631
Db 61 GGCCTTCGGAAGATCAATGTGCTTACTGCTGGAGATATGTCCTCCCATTTACAAT 120

QY 2632 AGTTTACTACAGTGGATGCTGCTGAGCCAGTGTGATGCTGGATTTCTGGAAATTTATCA 2691
Db 121 AGTTTACTACAGTGGATGCTGCTGAGCCAGTGTGATGCTGGATTTCTGGAAATTTATCA 180

QY 2692 GAACCCAGTACTCGGGAGTGAAGCTGCACTGCACTCAACCCACTCAATGGTAGTGC 2751
Db 181 GAACCCAGTACTCGGGAGTGAAGCTGCACTGCACTCAACCCACTCAATGGTAGTGC 240

QY 2752 TGTGAAGCCCTGCAACACACAGTGTCTAAGCAGTGCCTGAGCAGCACCATGTGCTTCGAGACA 2811
Db 241 TGTGAAGCCCTGCAACACACAGTGTCTAAGCAGTGCCTGAGCAGCACCATGTGCTTCGAGACA 300

QY 2812 GCATGTGAGATGTCACCAAGGCGAGCTCTGAGTGCATGTGGTGACGCAACATGAAGCAG 2871
Db 301 GCATGTGAGATGTCACCAAGGCGAGCTCTGAGTGCATGTGGTGACGCAACATGAAGCAG 360

QY 2872 TGTGAGGACTCAATGCTTGTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2931
Db 361 TGTGAGGACTCAATGCTTGTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 420

QY 2932 ACATGAGCAGCTGCCCCCTGAAAAATGTTTCAAGCTACTGTACCTGTAGTCAATGCTTG 2991
Db 421 ACATGAGCAGCTGCCCCCTGAAAAATGTTTCAAGCTACTGTACCTGTAGTCAATGCTTG 480

QY 2992 GAGCAACAGGCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3051
Db 481 GAGCAACAGGCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540

QY 3052 GAGGTTTCTTATAAGGACCACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3111
Db 541 GAGGTTTCTTATAAGGACCACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 3112 CCAGAGCCCTGCTCAATTCAGATGTGCTAGAGCAGCAGATCACTGCTGCTTTC 3171
Db 601 CCAGAGCCCTGCTCAATTCAGATGTGCTAGAGCAGCAGATCACTGCTGCTTTC 660

QY 3172 ATTCAGTGTCCAG 3184
Db 661 ATTCAGTGTCCAG 673

RESULT 15
CD644584
LOCUS AGENCOURT_14539338 NIA Human H1 Embryonic Stem Cell cDNA Library
DEFINITION (Long) Homo sapiens cDNA clone IMAGE:30421895 5', mRNA sequence.
ACCESSION CD644584
VERSION CD644584.1 GI:31816608
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 792)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
```

Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgaabs-x@mail.nih.gov  
Tissue Procurement: Irene Ginis and Mahendra Rao, NIA  
cDNA Library Preparation: Yulan Piao and Minoru KC  
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
Clone distribution: MGC c lone distribution information  
can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: NDAM500 row: 9 column: 24  
High quality sequence stop: 681.  
Location/Qualifiers  
1. .792  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30421895"  
/issue\_type="Embryonic Stem cells"  
/cell\_line="WA01"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIA Human H1 Embryonic Stem Cell cDNA Library (Long)"  
/note="Vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;  
This is a long-transcript enriched cDNA library (Genome  
Res. 11: 1553-1558 (2001)). [PMID: 11544199] from WA01  
cell line. Undifferentiated human ES cell line WA01/H1  
was obtained from Wicell Research Institute, Inc.,  
Madison, WI, cultured according to their instructions, on  
MEF feeders. They formed round colonies with defined edges  
and were positive for alkaline phosphatase, SSEA-4, OCT3,  
OCT4, REX1, UTR, TERT, SOX2, CX43 and CX45. They are  
negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLK3, SSEA-1,  
TUBB3, NES, GFAP, and EOMES. When confluent (18-10 days  
after plating), the ES cells from 4 X 6 cm dishes were  
treated with 1 mg/ml collagenase, type IV  
(Invitrogen/GIBCO) for 5-10 min and gently scraped off  
with 5 ml pipette. RNA was purified with Trizol Reagent  
from invitrogen. Protocol ref: Genome Res. 11: 1553-1558  
(2001). [PMID:11544199] Double-stranded cDNAs were  
synthesized with an oligo(dT) primer [Invitrogen:  
5'-pOACTAGTTCTAGATCGGCGCGCCCTTTTCTTTT-3'] from  
3.4g of total RNA, treated with T4 DNA polymerase, and  
purified by ethanol-precipitation. The cDNAs were ligated  
to Loner-linker LL-Sal4, purified by phenol/chloroform  
extraction, and separated from free linkers by  
Centricon-100 column. Then, the cDNAs were amplified by  
long-range high fidelity PCR using Ex Taq polymerase  
(Takara) with a primer Sal4-S for 25 cycles. The products  
were purified by phenol/chloroform extraction and  
Centricon-100 column. The cDNAs were digested with SalI  
and NotI enzymes and cloned into SalI/NotI site of  
pCMV-SPORT6 plasmid vector. The average insert size is  
about 3.6kb."

Query Match 15.5%; Score 663.8; DB 14; Length 792;  
Best Local Similarity 94.9%; Pred. No. 4.8e-116;  
Matches 686; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 479 GCGTCATTGAAGGACAGCCAAATAGATTAATGAGACTTCGTTTCAATCAATTTGCTACAG 538  
Db 12 GCGTCATTGAAGGACAGCCAAATAGATTAATGAGACTTCGTTTCAATCAATTTGCTACAG 71

QY 539 AGTGAGTGGGACCATTTATGATGGGACTCAATTTATGCGGCTGCTGCTGCTGCTGCTGCTG 598  
Db 72 AGTGAGTGGGACCATTTATGATGGGACTCAATTTATGCGGCTGCTGCTGCTGCTGCTGCTG 131

QY 599 CTGCAATTTAGTGGCTCAATTTCTCTGAGAGAGATGCAATGAGACTGTCCTGAGGTTG 658  
Db 132 CTGCAATTTAGTGGCTCAATTTCTCTGAGAGAGATGCAATGAGACTGTCCTGAGGTTG 191



Qy 659 TTGCCACATCAGGTTATGCTTCTGCTCAATTTTTTTAGTGTGCTGCTTATAATTTGACTG 718  
Db |||||||  
Qy 192 TTGCCACATCAGGTTATGCTTCTGCTCAATTTTTTTAGTGTGCTGCTTATAATTTGACTG 251  
Db |||||||  
Qy 719 GATTTAATATTACTTACAGTTTGTATATGTCTCAATAAAGTCTCAGCCGAGGAGAGT 778  
Db |||||||  
Qy 252 GATTTAATATTACTTACAGTTTGTATATGTCTCAATAAAGTCTCAGCCGAGGAGAGT 311  
Db |||||||  
Qy 779 GTAAGATCAGTATAGCAGGAACTGTTGAATGTGATGCTTCTGAAAAGTCTGAAAAGGTG 838  
Db |||||||  
Qy 312 GTAAGATCAGTATAGCAGGAACTGTTGAATGTGATGCTTCTGAAAAGTCTGAAAAGGTG 371  
Db |||||||  
Qy 839 AAGCATGTGACATTCCTCACTGTACAGACAACTGTGTTTTCCTCATCCAGGCACTGCA 898  
Db |||||||  
Qy 372 AAGCATGTGACATTCCTCACTGTACAGACAACTGTGTTTTCCTCATCCAGGCACTGCA 431  
Db |||||||  
Qy 899 ATTCAAGTGTGACAGGATGCTCTGCTTCTCAGACTGGCAGGCTCCTGGATGTTTCA 958  
Db |||||||  
Qy 432 ATTCAAGTGTGACAGGATGCTCTGCTTCTCAGACTGGCAGGCTCCTGGATGTTTCA 491  
Db |||||||  
Qy 959 TTCTGTACCAGCTAACAGTCAATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCC 1018  
Db |||||||  
Qy 492 TTCTGTACCAGCTAACAGTCAATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCC 551  
Db |||||||  
Qy 1019 CCAGAGCATCTCATAAAGCTGTGTCATATGGAACAATTTATGTGGTTTGTGGAGATATA 1078  
Db |||||||  
Qy 552 CCAGAGCATCTCATAAAGCTGTGTCATATGGAACAATTTATGTGGTTTGTGGAGATATA 611  
Db |||||||  
Qy 1079 TGTTCAACCACTCAGATTATAACATGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGC 1138  
Db |||||||  
Qy 612 TGTTCAACCACTCAGATTATAACATGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGC 671  
Db |||||||  
Qy 1139 TTCCACTAAACCGTTCTGTGAAACAATGTGTTGTAGATATGGTCAATTTTGGCAATTAT 1198  
Db |||||||  
Qy 672 TTCACATAACGTTCTGTGAACATGTGTTGTAGATAGGTCAATCTTTGGCATATACAGGAT 731  
Db |||||||  
Qy 1199 ACA 1201  
Db ||  
Qy 732 AAA 734  
Db ||

Search completed: March 5, 2004, 12:07:59  
Job time : 6801.26 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:06:34 ; Search time 1010.9 Seconds

(without alignments)  
18028.269 Million cell updates/sec

Title: US-09-787-097-13

Perfect score: 4290  
Sequence: 1 atggtggccagcagggcgc.....agcctgggacctgcacatctga 4290

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_29Jan04:.\*  
1: geneseqn1980s:.\*  
2: geneseqn1990s:.\*  
3: geneseqn2000s:.\*  
4: geneseqn2001as:.\*  
5: geneseqn2001bs:.\*  
6: geneseqn2002s:.\*  
7: geneseqn2003as:.\*  
8: geneseqn2003bs:.\*  
9: geneseqn2003cs:.\*  
10: geneseqn2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4290	100.0	4290	3	AAZ52276 Human mem
2	4290	100.0	8561	5	AAZ52276 DNA encod
3	4184.8	97.5	8590	3	AAZ91920 Human mah
4	3981	92.8	4068	3	AAZ52275 Human mem
5	3801.4	88.6	3819	3	AAZ52277 Human sol
6	3696.2	86.2	4072	3	AAZ91921 Human mah
7	3536	82.4	8739	9	ADBS3249 Primaty r
8	3469.4	80.9	3597	3	AAZ52274 Human sol
9	3469.4	80.9	3597	5	AAZ52274 DNA encod
10	3469.4	80.9	3597	9	AAZ52274 Human att
11	3396.4	79.2	8927	3	AAZ91916 Wild type
12	2807.2	65.4	3490	5	AAZ52260 DNA encod
13	2248.2	52.4	2625	3	AAZ91922 Human mah
14	1830.2	42.7	6373	3	AAZ91919 Human mah
15	1521.8	35.5	2419	3	AAZ91917 Murine ma
16	1479.4	34.5	4134	7	ABZ58710 Murine ho
17	1467	34.2	4140	7	ABZ58709 Human hom
18	1133.4	26.4	3609	6	ABZ58709 Human cDN
19	873.2	20.4	3165	5	AAZ58781 DNA encod
20	873.2	20.4	3313	5	AAZ58781 DNA encod
21	819.4	19.1	1012	2	AAZ58781 DNA encod
22	812.8	18.9	6517	4	AAH98748 Human EST
23	773.2	18.0	6532	5	AAZ586766 DNA encod

24	640.2	14.9	1051	3	AAZ91918 Murine ma
25	420	9.8	1654	7	ABZ34730 Human mdd
26	415.2	9.7	1383	6	AAZ54064 CDNA of h
27	404.2	9.4	671	5	AAZ72658 DNA encod
28	314.4	7.3	668	4	AAZ02409 Human sec
29	277.4	6.5	481	4	AAZ14686 Human bre
30	276.4	6.4	292	3	AAZ45505 Human sec
31	267.6	6.2	845	6	ABQ52394 Oligonuc1
32	267.6	6.2	845	6	ABQ52395 Oligonuc1
33	266.4	6.2	482	4	AAZ23548 Human bre
34	253.8	5.9	982	6	ABZ39784 Human NS
35	249.2	5.8	531	4	AAZ43425 Probe #12
36	249.2	5.8	531	4	AAZ11523 Human gen
37	246	5.7	246	6	ABK38928 CDNA enco
38	246	5.7	246	6	ABK38928 Human lun
39	246	5.7	246	7	ACA02443 Lung "canc
40	236	5.5	657	7	ABZ58714 Upregulat
41	234.4	5.5	207433	5	ABZ72040 Gene 216
42	234.4	5.5	207433	7	ABZ74891 BAC109812
43	209.8	4.9	5312	6	ABL32196 Human imm
44	206	4.8	845	6	ABQ52396 Oligonuc1
45	206	4.8	845	6	ABQ52397 Oligonuc1

ALIGNMENTS

RESULT 1  
AAZ52276  
ID AAZ52276 standard; CDNA; 4290 BP.  
XX AC AAZ52276;  
DT DT 18-JUL-2000 (first entry)  
XX DE Human membrane attractin-2 cDNA.  
XX KW Human; membrane attractin-2; immune response; macrophage; monocyte;  
XX KW T cell; immunostimulant; immunosuppressed patient; cancer;  
XX KW immunodeficiency syndrome; transplant; autoimmune disease; ss.  
XX OS Homo sapiens.  
XX PH Key Location/Qualifiers  
XX CDS 1..4290  
FT FT /\*tag= a  
FT FT /product= "Membrane attractin-2"  
XX EN WO200015651-A1.  
XX PD 23-MAR-2000.  
XX PF 14-SEP-1999; 99WO-US020948.  
XX PR 14-SEP-1998; 98US-0100137P.  
XX PA (DAND ) DANA FARBER CANCER INST INC.  
XX PI Duke-Cohan JS, Schlossman SP;  
XX DR WPI; 2000-271373/23.  
XX DR P-PSDB; AAY70691.  
XX PT Isolated nucleic acids encoding human attractin polypeptides useful for  
XX PT enhancing immune responses.  
XX PS Claim 3; Fig 14; 120pp; English.  
XX CC The patent discloses four forms of human attractin polypeptides which  
XX CC enhance immune response by promoting macrophage and monocyte spreading in  
XX CC the presence of T cells. These include soluble attractin-1 and -2 and  
XX CC membrane attractin-1 and -2. These various forms of attractin are encoded  
XX CC by alternatively spliced mRNA molecule transcribed from a single gene.



QY 1921 CTGGTATTCACCTCGGAAACAGTGTGANGCGCATCGGAGTGAAGCCGCTTGTTTACAGCA 1980  
Db 1921 CTGGTATTCACCTCGGAAACAGTGTGANGCGCATCGGAGTGAAGCCGCTTGTTTACAGCA 1980  
QY 1981 GGAACCTGGTATTCGGTGTGTGGAAACACAGAGTGTCTCAGTGTATCTCGTGGGCGGTG 2040  
Db 1981 GGAACCTGGTATTCGGTGTGTGGAAACACAGAGTGTCTCAGTGTATCTCGTGGGCGGTG 2040  
QY 2041 GCAACTGATGAACAAGAAAGTTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGAC 2100  
Db 2041 GCAACTGATGAACAAGAAAGTTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGAC 2100  
QY 2101 CATGACAGATGTACACAGACACAGATGTGTACAGCTGTACAGCCAAACCAATGACTGC 2160  
Db 2101 CATGACAGATGTACACAGACACAGATGTGTACAGCTGTACAGCCAAACCAATGACTGC 2160  
QY 2161 CACTGSGCAATGACCAATGTGTGCCAGGAACACACAGCTGTCTCAGAGGCCAGATCTCC 2220  
Db 2161 CACTGSGCAATGACCAATGTGTGCCAGGAACACACAGCTGTCTCAGAGGCCAGATCTCC 2220  
QY 2221 ATTTTATGATGAGAAATGCCCCCAGGATAACCCCATGTACTACTGTACACAGAGACC 2280  
Db 2221 ATTTTATGATGAGAAATGCCCCCAGGATAACCCCATGTACTACTGTACACAGAGACC 2280  
QY 2281 AGCTGAGGAGCTGTGCCCTGGACACAGAACTGCCAGTGGAGCCCGGAAATCAGGAGTGC 2340  
Db 2281 AGCTGAGGAGCTGTGCCCTGGACACAGAACTGCCAGTGGAGCCCGGAAATCAGGAGTGC 2340  
QY 2341 ATTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTGGAACTCATGTTTG 2400  
Db 2341 ATTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTGGAACTCATGTTTG 2400  
QY 2401 AAAATTTACTCTGCCAGGAGAAATATGACAACTGTAATTTGTTCTGTAGGAACCAAT 2460  
Db 2401 AAAATTTACTCTGCCAGGAGAAATATGACAACTGTAATTTGTTCTGTAGGAACCAAT 2460  
QY 2461 GCCCTTTTGGCTTCTTTACAAACCCAGAAAGGTAGAATTTGTCCTTAAGCAGTGGGA 2520  
Db 2461 GCCCTTTTGGCTTCTTTACAAACCCAGAAAGGTAGAATTTGTCCTTAAGCAGTGGGA 2520  
QY 2521 ATATGAGTCACTCAGAGCATGTCCAAGCTCACTTACCCCATGGGTGGCTTCGG 2580  
Db 2521 ATATGAGTCACTCAGAGCATGTCCAAGCTCACTTACCCCATGGGTGGCTTCGG 2580  
QY 2581 AAGATCAATGTCTACTGTCTGGGAAAGATGTGCCCATTTACAAATAGTTTACTA 2640  
Db 2581 AAGATCAATGTCTACTGTCTGGGAAAGATGTGCCCATTTACAAATAGTTTACTA 2640  
QY 2641 CAGTGGATGCGCTCAGGCCAGTGTGCTGGATTTCTGTGGAAATTTATCAGAACCCAGT 2700  
Db 2641 CAGTGGATGCGCTCAGGCCAGTGTGCTGGATTTCTGTGGAAATTTATCAGAACCCAGT 2700  
QY 2701 ACTGGGAGCTGAAGGCTGCAACCTGCATCAACCACTCAATGGTGTGTCTGTGAAGG 2760  
Db 2701 ACTGGGAGCTGAAGGCTGCAACCTGCATCAACCACTCAATGGTGTGTCTGTGAAGG 2760  
QY 2761 CCTGCAAAACCAAGTGTCTAAGCAGTGGCGGACACCAATGTGCTTGTAGGACAGCATGTGA 2820  
Db 2761 CCTGCAAAACCAAGTGTCTAAGCAGTGGCGGACACCAATGTGCTTGTAGGACAGCATGTGA 2820  
QY 2821 GATTGACACGCGAGCTCTGAGTGCATGTGTGTGACACACATGAGCAGTGTGTGGAC 2880  
Db 2821 GATTGACACGCGAGCTCTGAGTGCATGTGTGTGACACACATGAGCAGTGTGTGGAC 2880  
QY 2881 TCCAAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940  
Db 2881 TCCAAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940  
QY 2941 ACCTGCCCCCTGAAATTTGTTTCCAGCTACTGTACTGTCTGAGGCTCCTTGGAGCAACCA 3000  
Db 2941 ACCTGCCCCCTGAAATTTGTTTCCAGCTACTGTACTGTCTGAGGCTCCTTGGAGCAACCA 3000

QY 3001 GGCTGTGGCTGTCTACTGATCCAGCAATACTGGCAAGGAAATGCATAGAGGGTTCC 3060  
Db 3001 GGCTGTGGCTGTCTACTGATCCAGCAATACTGGCAAGGAAATGCATAGAGGGTTCC 3060  
QY 3061 TATAAAGAACAGTGAAGATGCTTCGGAAGCCCTACAGGAAATTTCTATCCAGAGCC 3120  
Db 3061 TATAAAGAACAGTGAAGATGCTTCGGAAGCCCTACAGGAAATTTCTATCCAGAGCC 3120  
QY 3121 CTGCTCAATTCACATGTGTCTAGAGACAGAGATACAACTGCTCTTTCAATCACTGT 3180  
Db 3121 CTGCTCAATTCACATGTGTCTAGAGACAGAGATACAACTGCTCTTTCAATCACTGT 3180  
QY 3181 CCAGCTTGGCAATGCAACGGCCACAGTAATGCAATCAGAGCATCTGTGAGAAGTGT 3240  
Db 3181 CCAGCTTGGCAATGCAACGGCCACAGTAATGCAATCAGAGCATCTGTGAGAAGTGT 3240  
QY 3241 GAGAACCTGACACAGCAAGCACTGCGAGACCTGCAATCTGCTTCTACGCTGATCCC 3300  
Db 3241 GAGAACCTGACACAGCAAGCACTGCGAGACCTGCAATCTGCTTCTACGCTGATCCC 3300  
QY 3301 ACCAATGAGGAGAAATGTCAAGTGAATGGCAACGCTCTCTGTGCAACACC 3360  
Db 3301 ACCAATGAGGAGAAATGTCAAGTGAATGGCAACGCTCTCTGTGCAACACC 3360  
QY 3361 AACCGGCAAGTGTCTGCAACCAAGGGCGTCAAGGGGACGAGTGCAGCTATGT 3420  
Db 3361 AACCGGCAAGTGTCTGCAACCAAGGGCGTCAAGGGGACGAGTGCAGCTATGT 3420  
QY 3421 GAGGTAGAAATTCGATACCAAGGAAACCTCTCAGAGGAAATGTTATTAATCTTCTT 3480  
Db 3421 GAGGTAGAAATTCGATACCAAGGAAACCTCTCAGAGGAAATGTTATTAATCTTCTT 3480  
QY 3481 ATGTACTATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTATTAACAGCTATC 3540  
Db 3481 ATGTACTATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTATTAACAGCTATC 3540  
QY 3541 AATTTTGGCTTACTCTTCGAGCAACAAAGGGATTTGGACATGTTCAATGCCCTCC 3600  
Db 3541 AATTTTGGCTTACTCTTCGAGCAACAAAGGGATTTGGACATGTTCAATGCCCTCC 3600  
QY 3601 AAGAAATTTCAACCTCAACATCACCTGGCTGCCAGTTCCTCAGCTGGAACCCAGGCTGGA 3660  
Db 3601 AAGAAATTTCAACCTCAACATCACCTGGCTGCCAGTTCCTCAGCTGGAACCCAGGCTGGA 3660  
QY 3661 GAAGAGATGCTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720  
Db 3661 GAAGAGATGCTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720  
QY 3721 GAGAAATTTGATTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTCAAGTAATTC 3780  
Db 3721 GAGAAATTTGATTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTCAAGTAATTC 3780  
QY 3781 ACTGSCCATCAAAATTCAGATTCCTCTCAGCAGCAAGCAATTTATGACCTGGA 3840  
Db 3781 ACTGSCCATCAAAATTCAGATTCCTCTCAGCAGCAAGCAATTTATGACCTGGA 3840  
QY 3841 CAGTCTTCGCTCAGTTCCTTCAGTGTGTTTCCCTCTTTTGTCTCCTGTGGCTCTGTGGTT 3900  
Db 3841 CAGTCTTCGCTCAGTTCCTTCAGTGTGTTTCCCTCTTTTGTCTCCTGTGGCTCTGTGGTT 3900  
QY 3901 TGAAGATCAAAACAAAGTTGTTGGGCTCCAGACGTAGAGCAACTCTTCTGAGAGATG 3960  
Db 3901 TGAAGATCAAAACAAAGTTGTTGGGCTCCAGACGTAGAGCAACTCTTCTGAGAGATG 3960  
QY 3961 CAACAGATGGCAGCGCTTCCCTCTGCTGTAAATGTCCTTGGAAACAGATGAGGAG 4020  
Db 3961 CAACAGATGGCAGCGCTTCCCTCTGCTGTAAATGTCCTTGGAAACAGATGAGGAG 4020  
QY 4021 CCTCTCTGATCTTATTTGGGGGAGTAAAGACTGTTCCAAACCCATTCGACTGGAGCG 4080  
Db 4021 CCTCTCTGATCTTATTTGGGGGAGTAAAGACTGTTCCAAACCCATTCGACTGGAGCG 4080  
QY 4081 TGTTTTGGCAACAAAGCCGCTGCTCTCTGTGTTGTGAGGCTCCTTCGAGGCTGGT 4140

Db 4081 TGTATTGGCAACAAAGCGCTGCTCTCTCTGTGTTGTGAGGCTCCCTCGAGGCGCTGGGT 4140  
QY 4141 GGCATCCCTCCTCTGGGCACTCAGGCTTGTCTGTGCTGGCCAGCGCCCTGGTGGACATTTCT 4200  
Db 4141 GGCATCCCTCCTCTGGGCACTCAGGCTTGTCTGTGCTGGCCAGCGCCCTGGTGGACATTTCT 4200  
QY 4201 CAGCAGATGCCGATAGTGTACAAGGAGAGAGTCCAGGAGCCGTGAGAAACCGGAGCAGCAG 4260  
Db 4201 CAGCAGATGCCGATAGTGTACAAGGAGAGTCCAGGAGCCGTGAGAAACCGGAGCAGCAG 4260  
QY 4261 CCCCTGACAGCCTGGGACCTGCATCTGA 4290  
Db 4261 CCCCTGACAGCCTGGGACCTGCATCTGA 4290

## RESULT 2

AA572659  
ID AA572659 standard; cDNA; 8561 BP.

XX  
AC AA572659;

XX  
DT 13-FEB-2002 (first entry)

XX  
DE DNA encoding novel human diagnostic protein #8463.

XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX  
OS Homo sapiens.

XX  
PN WO200175067-A2.

XX  
PD 11-OCT-2001.

XX  
PF 30-MAR-2001; 2001WO-US008631.

XX  
PR 31-MAR-2000; 2000US-00540217.

XX  
PX 23-AUG-2000; 2000US-00649167.

XX  
FA (HYSE-) HYSEQ INC.

XX  
FI Drmanac RT, Liu C, Tang YT;

XX  
DR WPI: 2001-639362/73.

XX  
P-PSDB; ABG08472.

XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX  
PS Claim 1; SEQ ID NO 8463; 103pp; English.

XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful for medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AA564197-AA594564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in

CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 8561 BP; 2246 A; 1983 C; 2065 G; 2267 T; 0 U; 0 Other;  
Query Match 100.0%; Score 4290; DB 5; Length 8561;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGTGGCGGCGAGCGGCGGCACTGAGGCAAGGCTGAGGAGGAGGAGCGGCGGCGGCA 60  
Db 1 ATGGTGGCGGCGAGCGGCGGCACTGAGGCAAGGCTGAGGAGGAGGAGCGGCGGCGGCA 60  
QY 61 GCGCTCGGCGGCGAGGAGCGGCGGCGGCGGCACTGGGACTGGGAGCTGAGGAGGAGGAGG 120  
Db 61 GCGCTCGGCGGCGAGGAGCGGCGGCGGCGGCACTGGGACTGGGAGCTGAGGAGGAGGAGG 120  
QY 121 CCGGGGCTGGGGGCGGCGGCTGCGCTCCCGCGGCTGTGTCTCCACCGCTGGGCGCAGG 180  
Db 121 CCGGGGCTGGGGGCGGCGGCTGCGCTCCCGCGGCTGTGTCTCCACCGCTGGGCGCAGG 180  
QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAG 240  
Db 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGAG 240  
QY 241 GCGGAGGCGGCGGCGGCGGCGGCGGCTGCGGCTCAGCGCGGCGGCGGCGGCGGCGGCGG 300  
Db 241 GCGGAGGCGGCGGCGGCGGCGGCGGCGGCTGCGGCTCAGCGCGGCGGCGGCGGCGGCGG 300  
QY 301 TGTGACCGGCGGCTGCTCAACGGGCGGCTGCTGCAACCTGCGGCGGCGGCGGCGGCGGCTGC 360  
Db 301 TGTGACCGGCGGCTGCTCAACGGGCGGCTGCTGCAACCTGCGGCGGCGGCGGCGGCGGCTGC 360  
QY 361 CCGGCGGCTGGGCGGCGGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420  
Db 361 CCGGCGGCTGGGCGGCGGCGGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420  
QY 421 TCTTCTGGGTTGTGACAGATGGGAACTTATAAATACAAAGAGAGTGGCGGCGGCGGCGG 480  
Db 421 TCTTCTGGGTTGTGACAGATGGGAACTTATAAATACAAAGAGAGTGGCGGCGGCGGCGG 480  
QY 481 CTCAATTGAAGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540  
Db 481 CTCAATTGAAGACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540  
QY 541 TGTAGTTGGGACCAATTATATGTTTATGATGGGAGCTCAATTTATGACCGCTAGTTGCT 600  
Db 541 TGTAGTTGGGACCAATTATATGTTTATGATGGGAGCTCAATTTATGACCGCTAGTTGCT 600  
QY 601 GCATTTAGTGGCGCTCATTTGTTTCTGAGAGAGATGGCAATGAGACTGCTGAGGTTGTT 660  
Db 601 GCATTTAGTGGCGCTCATTTGTTTCTGAGAGAGATGGCAATGAGACTGCTGAGGTTGTT 660  
QY 661 GCCACATCAGGTTATGCTTGTGCTGCTCAATTTTATGATGCTGCTTATAATTTGACTGGA 720  
Db 661 GCCACATCAGGTTATGCTTGTGCTGCTCAATTTTATGATGCTGCTTATAATTTGACTGGA 720  
QY 721 TTTAATATTACTACAGTTTGTATGTTTGTCCCAATTAATGCTCAGGCGGAGAGTGT 780  
Db 721 TTTAATATTACTACAGTTTGTATGTTTGTCCCAATTAATGCTCAGGCGGAGAGTGT 780  
QY 781 AAGATCAGTAAATAGCAGCGGAACTGTTGAATGTGAATGTTTCTGAAAGCTGAAAGGTGAA 840  
Db 781 AAGATCAGTAAATAGCAGCGGAACTGTTGAATGTGAATGTTTCTGAAAGCTGAAAGGTGAA 840  
QY 841 GCATGTGACATTCCTCTACTGTACAGCAACTGTGGTTTCTCTATCGAGGCACTGCAAT 900  
Db 841 GCATGTGACATTCCTCTACTGTACAGCAACTGTGGTTTCTCTATCGAGGCACTGCAAT 900  
QY 901 TCAAGTGTATGTACAGGAGTGTCTCTGCTTCTCAGACTGGGCGGCTCTGATGTTTCAAGTT 960  
Db 901 TCAAGTGTATGTACAGGAGTGTCTCTGCTTCTCAGACTGGGCGGCTCTGATGTTTCAAGTT 960

QY 961 CCTGTACAGCTACCAAGTCAATTTTGGACTCGAGAGGAATATTTCTAACTTAAAGCTCCCC 1020  
Db 961 CCTGTACAGCTACCAAGTCAATTTTGGACTCGAGAGGAATATTTCTAACTTAAAGCTCCCC 1020  
QY 1021 AGAGCATCTCATAAAGCTGTGGTCAATGAGAAACATTAATGAGGTTGTGTGAGGATATATG 1080  
Db 1021 AGAGCATCTCATAAAGCTGTGGTCAATGAGAAACATTAATGAGGTTGTGTGAGGATATATG 1080  
QY 1081 TTCAACCACTCAGATTATACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140  
Db 1081 TTCAACCACTCAGATTATACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140  
QY 1141 CCACATAACCGTCTGTGAAACATGCTGTTGTAGATPATGGTCAATCTTTTGGCAATTATAC 1200  
Db 1141 CCACATAACCGTCTGTGAAACATGCTGTTGTAGATPATGGTCAATCTTTTGGCAATTATAC 1200  
QY 1201 AAGGATAAATTTACATGTTATGAGGAAATTTGATCCAACTGGGATGTGACCAATGAG 1260  
Db 1201 AAGGATAAATTTACATGTTATGAGGAAATTTGATCCAACTGGGATGTGACCAATGAG 1260  
QY 1261 TTGAGAGTTTTCACATTCATAATGAGTCAATGGGTGTTGTGACCCCTAAGGCAAAAGGAG 1320  
Db 1261 TTGAGAGTTTTCACATTCATAATGAGTCAATGGGTGTTGTGACCCCTAAGGCAAAAGGAG 1320  
QY 1321 CAGTATGAGTGGTGGGCACTCTGCACACATTTGTTACCTGAAAGATGGCCGAGTGGTC 1380  
Db 1321 CAGTATGAGTGGTGGGCACTCTGCACACATTTGTTACCTGAAAGATGGCCGAGTGGTC 1380  
QY 1381 ATGCTGGTCACTTTTGGTCACTGCCCCTCTCATGAGATATATAGCAATGTGCAGGAATAT 1440  
Db 1381 ATGCTGGTCACTTTTGGTCACTGCCCCTCTCATGAGATATATAGCAATGTGCAGGAATAT 1440  
QY 1441 GATTGGATAAGAACACATGAGTATATACACACCCAGGGTGGCTTGTGCAAGGGGGT 1500  
Db 1441 GATTGGATAAGAACACATGAGTATATACACACCCAGGGTGGCTTGTGCAAGGGGGT 1500  
QY 1501 TACGGCCATACAGTGTGTAGCATAGGACAGGAGGCTTATAGCTTCACTGTGGCTAC 1560  
Db 1501 TACGGCCATACAGTGTGTAGCATAGGACAGGAGGCTTATAGCTTCACTGTGGCTAC 1560  
QY 1561 AAGGCTTTCAGTGCCAAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGGAT 1620  
Db 1561 AAGGCTTTCAGTGCCAAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGGAT 1620  
QY 1621 ACCGATGTGGACCATCTTAAAGACAGCCGATTTTCGTTACTTGTGCACAGCTGTG 1680  
Db 1621 ACCGATGTGGACCATCTTAAAGACAGCCGATTTTCGTTACTTGTGCACAGCTGTG 1680  
QY 1681 ATAGTGTGTGAAACCATGCTGTGTGTTGGGGAAACACACACATGACATCTATGAGC 1740  
Db 1681 ATAGTGTGTGAAACCATGCTGTGTGTTGGGGAAACACACACATGACATCTATGAGC 1740  
QY 1741 CATGGCGCAATGCTTCTTTCAGATTTCAATGCGCTATGACATGCTGTGACCGCTGG 1800  
Db 1741 CATGGCGCAATGCTTCTTTCAGATTTCAATGCGCTATGACATGCTGTGACCGCTGG 1800  
QY 1801 TCAGTGTCTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860  
Db 1801 TCAGTGTCTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860  
QY 1861 TTACACACAGCAGCATGATGTGTGCTGGTGTTCATAGTCTCTCTCCCTCAGCGACATC 1920  
Db 1861 TTACACACAGCAGCATGATGTGTGCTGGTGTTCATAGTCTCTCTCTCAGCGACATC 1920  
QY 1921 CTGGTATTTCACTCGGAACAGTGTGATGCGGATCGAGTGAAGCGGCTTGTATGACGCA 1980  
Db 1921 CTGGTATTTCACTCGGAACAGTGTGATGCGGATCGAGTGAAGCGGCTTGTATGACGCA 1980  
QY 1981 GGACCTGGTATTCGGTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTCGTGGCGCTG 2040  
Db 1981 GGACCTGGTATTCGGTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTCGTGGCGCTG 2040  
QY 2041 GCAACTGATGAACAGAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTCTTGAC 2100

Db 2041 GCAACTGATGAACAGAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTCTTGAC 2100  
QY 2101 CATGACAGATGACACAGACACAGATTGTACAGCTGTACAGCCAAACACCAATGACTGC 2160  
Db 2101 CATGACAGATGACACAGACACAGATTGTACAGCTGTACAGCCAAACACCAATGACTGC 2160  
QY 2161 CACTGGTCAATGACCATTTGTCTCCAGGAACACACAGCTGTCTAGAGGCGCAGATCTCC 2220  
Db 2161 CACTGGTCAATGACCATTTGTCTCCAGGAACACACAGCTGTCTAGAGGCGCAGATCTCC 2220  
QY 2221 ATTTTGTAGTATGAGATTTGCCCAAGGATACCCCATGTACTCTGTAAACCAAGAGAC 2280  
Db 2221 ATTTTGTAGTATGAGATTTGCCCAAGGATACCCCATGTACTCTGTAAACCAAGAGAC 2280  
QY 2281 AGCTGACAGAGCTGTGCTCCCTGACACAGAACTGCCAGTGGGAGCCCGGAATCAGAGTGC 2340  
Db 2281 AGCTGACAGAGCTGTGCTCCCTGACACAGAACTGCCAGTGGGAGCCCGGAATCAGAGTGC 2340  
QY 2341 ATTGCCCTCCCGAAAAATATCTGTGGCATTTGCTGGCATTTGGTTGGAAACTCATGTTTG 2400  
Db 2341 ATTGCCCTCCCGAAAAATATCTGTGGCATTTGCTGGCATTTGGTTGGAAACTCATGTTTG 2400  
QY 2401 AAAATTTACTTACGCAAGGAGAAATATGACAAATGCTAAATTTCTGTAGGAACCACAAT 2460  
Db 2401 AAAATTTACTTACGCAAGGAGAAATATGACAAATGCTAAATTTCTGTAGGAACCACAAT 2460  
QY 2461 GGCCTTTTGGCTTCTTAAACCCAGAAAGGATAGAAATTTGCTTTAAGCAGCTGGA 2520  
Db 2461 GGCCTTTTGGCTTCTTAAACCCAGAAAGGATAGAAATTTGCTTTAAGCAGCTGGA 2520  
QY 2521 ATAAAGCACTCATCTCAGAGCATGTCACAGCTCACTTAAACCCATGGCTGGCTTGG 2580  
Db 2521 ATAAAGCACTCATCTCAGAGCATGTCACAGCTCACTTAAACCCATGGCTGGCTTGG 2580  
QY 2581 AAGATCAATGTGTCTTACTTCTGGGAGGATATGTCCCAATTTACAAATAGTTTACTA 2640  
Db 2581 AAGATCAATGTGTCTTACTTCTGGGAGGATATATGTCCCAATTTACAAATAGTTTACTA 2640  
QY 2641 CAGTGTAGCTGTGAGCCAGTGTGATGCTGGAATTTCTGGAATTTATCAGAACCCAGT 2700  
Db 2641 CAGTGTAGCTGTGAGCCAGTGTGATGCTGGAATTTCTGGAATTTATCAGAACCCAGT 2700  
QY 2701 ACTCGGGACTGAAGGCTCAACCTGCATCAACCCACTCAATGGTAGTCTGTGGAAGG 2760  
Db 2701 ACTCGGGACTGAAGGCTCAACCTGCATCAACCCACTCAATGGTAGTCTGTGGAAGG 2760  
QY 2761 CTTGCAAAACACAGTGTCTAAGCAGTCCCGGACACATGTGCTTGGGACAGCATGTGA 2820  
Db 2761 CTTGCAAAACACAGTGTCTAAGCAGTCCCGGACACATGTGCTTGGGACAGCATGTGA 2820  
QY 2821 GATTGACACAGCGGAGCTGTGATGTCATGTGGTGAGCAACATGAAGCAGTGTGTGGAC 2880  
Db 2821 GATTGACACAGCGGAGCTGTGATGTCATGTGGTGAGCAACATGAAGCAGTGTGTGGAC 2880  
QY 2881 TCCAATGCTTATGTGGCTCTTCTCCCTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940  
Db 2881 TCCAATGCTTATGTGGCTCTTCTCCCTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940  
QY 2941 ACCTGCCCCCTGAAAAATTTGTTTCAGGCTACTGTACCTGTAGTCAATTTGGAGCAACCA 3000  
Db 2941 ACCTGCCCCCTGAAAAATTTGTTTCAGGCTACTGTACCTGTAGTCAATTTGGAGCAACCA 3000  
QY 3001 GGCTGTGGCTGTGTACTGTATCCAGCAATACTTGGCAAAAGGGAATGCAATAGAGGTTCC 3060  
Db 3001 GGCTGTGGCTGTGTACTGTATCCAGCAATACTTGGCAAAAGGGAATGCAATAGAGGTTCC 3060  
QY 3061 TATAAGGACCACTGAGATGCTTCCGAAAGGCTTACAGGAAATTTCTATCCAGAGCC 3120  
Db 3061 TATAAGGACCACTGAGATGCTTCCGAAAGGCTTACAGGAAATTTCTATCCAGAGCC 3120  
QY 3121 CTGCTCAATTTCCAGCATGTCTGTAGAGGACAGCAGATACAACTGGTCTTCTCATCTACTGT 3180



3121	DB	CTGCTCAATTCCAGCATGTGTCCTAGAGCAGCAGAGATCAAACTGGTCTTTTCATTCACTGT	3150
3181	QY	CCAGCTTGCCAAATGCAACGGGCCACAGTAAATGCCATCAATCAGAGCATCTGTGAGAAGTGT	3240
3181	DB	CCAGCTTGCCAAATGCAACGGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT	3240
3241	QY	GAGAACCTGACACAGGCGAAGCCTGCGAGACCTGCAATATCTGGCTTCTACGGTGATCCC	3300
3241	DB	GAGAACCTGACACAGGCGAAGCCTGCGAGACCTGCAATATCTGGCTTCTACGGTGATCCC	3300
3301	QY	ACCAATGGAGGAAATGTCAAGTGCAAATGGCAGCGGTCTCTGTGCAACACC	3360
3301	DB	ACCAATGGAGGAAATGTCAAGTGCAAATGGCAGCGGTCTCTGTGCAACACC	3360
3361	QY	AACACGGGCAATGCTTCTGCACCAACAAAGGCGCTAAGGGGACGAGTGCACGTATGT	3420
3361	DB	AACACGGGCAATGCTTCTGCACCAACAAAGGCGCTAAGGGGACGAGTGCACGTATGT	3420
3421	QY	GAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATATACCTTCTT	3480
3421	DB	GAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATATACCTTCTT	3480
3481	QY	ATTGACTATCAGTTTCACTTTAGTCTTATCCAGGAAGATGATCGCTATTTACACGCTATC	3540
3481	DB	ATTGACTATCAGTTTCACTTTAGTCTTATCCAGGAAGATGATCGCTATTTACACGCTATC	3540
3541	QY	AATTTTCTGGCTACTCTCTGACGACAAACAGGGATTTGGACATGTTTCATCAATGCCTCC	3600
3541	DB	AATTTTCTGGCTACTCTCTGACGACAAACAGGGATTTGGACATGTTTCATCAATGCCTCC	3600
3601	QY	AAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGCAACCCAGGCTGGA	3660
3601	DB	AAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGCAACCCAGGCTGGA	3660
3661	QY	GAGAGATGCTGTGTTTCAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTTAAT	3720
3661	DB	GAGAGATGCTGTGTTTCAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTTAAT	3720
3721	QY	GAGAAGTTTGATTTTCGCAACCCCAAAATATCACTTTCTTTGTTATGTTCAGTAATTTTC	3780
3721	DB	GAGAAGTTTGATTTTCGCAACCCCAAAATATCACTTTCTTTGTTATGTTCAGTAATTTTC	3780
3781	QY	ACTTGGCCCATCAAAATTCAGATTCGCTTCTCTCAGCACAGCAATTTTATGGAACCTGGTA	3840
3781	DB	ACTTGGCCCATCAAAATTCAGATTCGCTTCTCTCAGCACAGCAATTTTATGGAACCTGGTA	3840
3841	QY	CAGTTCTTCGTGACTTCTTCAGTTGTTTCCCTCTCTTTGCTCTGCTGGCTGTGGTT	3900
3841	DB	CAGTTCTTCGTGACTTCTTCAGTTGTTTCCCTCTCTTTGCTCTGCTGGCTGTGGTT	3900
3901	QY	TGGAAGATCAAAACAAAGTTGTGGGCTCCAGACGTAGAGCAACCTTCTTCGAGAGATG	3960
3901	DB	TGGAAGATCAAAACAAAGTTGTGGGCTCCAGACGTAGAGCAACCTTCTTCGAGAGATG	3960
3961	QY	CACAGATGCGACCGTCCCTTTGCTCTGTAATGTGCGCTTGGAAACAGATGAGGAG	4020
3961	DB	CACAGATGCGACCGTCCCTTTGCTCTGTAATGTGCGCTTGGAAACAGATGAGGAG	4020
4021	QY	CCTCCTGATCTTATTTGGGGGAGTATAAAGACTGTTCCTCCAAACCCATTTGACCTGGAGCCG	4080
4021	DB	CCTCCTGATCTTATTTGGGGGAGTATAAAGACTGTTCCTCCAAACCCATTTGACCTGGAGCCG	4080
4081	QY	TGTTTTGGCAACAAAGCCGTGCTCTCTGTGTTGTGAGGCTCCCTCGAGGCCCTGGGT	4140
4081	DB	TGTTTTGGCAACAAAGCCGTGCTCTCTGTGTTGTGAGGCTCCCTCGAGGCCCTGGGT	4140
4141	QY	GGCATCCCTCCTCTCTGGGCAGTCAGGTCTTGTGTCGGCCAGCGCCCTGGTGACATTTCT	4200
4141	DB	GGCATCCCTCCTCTCTGGGCAGTCAGGTCTTGTGTCGGCCAGCGCCCTGGTGACATTTCT	4200
4201	QY	CACGAGTCCGATAGTGTTACAAAGGAAAGTCAAGAGCCGTGAGAAACCCGAAACGACAG	4260
4201	DB	CACGAGTCCGATAGTGTTACAAAGGAAAGTCAAGAGCCGTGAGAAACCCGAAACGACAG	4260



Db 2220 TTTTAGGTATGAGAAATGCCCCAGGATAAACCCTATGTACTGTAAACAAGAACCCAG 2279  
Qy 2283 CTGCGAGGAGCTGTGCCCTGGACCAACTGCCAGTGGGAGCCCGGGAATCAGGAGTGCAT 2342  
Db 2280 CTGCGAGGAGCTGTGCCCTGGACCAACTGCCAGTGGGAGCCCGGGAATCAGGAGTGCAT 2339  
Qy 2343 TCCCTTGCCGCAAAATATCTGTGGCAATTTGGCTGGCAATTTGGTGTGAAACTCATGTTTGA 2402  
Db 2340 TCCCTTGCCGCAAAATATCTGTGGCAATTTGGCTGGCAATTTGGTGTGAAACTCATGTTTGA 2399  
Qy 2403 AATTACTATGCAAGGAGGAATATGCAATGCTTAATTTGCTGTAGGACCAATGTC 2462  
Db 2400 AATTACTATGCAAGGAGGAATATGCAATGCTTAATTTGCTGTAGGACCAATGTC 2459  
Qy 2463 CTTTTTGGCTTCTCTTACAAACCAGGAAGGTAGAAATTTGCTCTTAAGCAGCTGCGAAT 2522  
Db 2460 CTTTTTGGCTTCTCTTACAAACCAGGAAGGTAGAAATTTGCTCTTAAGCAGCTGCGAAT 2519  
Qy 2523 AATGAGATCATCTCAGAGCATGTCCAGCTCACCCTTAACCCCATGGGTGGCCCTTCGGAA 2582  
Db 2520 AATGAGATCATCTCAGAGCATGTCCAGCTCACCCTTAACCCCATGGGTGGCCCTTCGGAA 2579  
Qy 2583 GATCAATGTGCTTACTGCTGCGAAGATATGTCCCAATTTACAAATAGTTTACTACA 2642  
Db 2580 GATCAATGTGCTTACTGCTGCGAAGATATGTCCCAATTTACAAATAGTTTACTACA 2639  
Qy 2643 GTGGATGCCGTCTGAGCCAGTGTGGATTTCTGTGGAATTTATCAGAACCCAGTAC 2702  
Db 2640 GTGGATGCCGTCTGAGCCAGTGTGGATTTCTGTGGAATTTATCAGAACCCAGTAC 2699  
Qy 2703 TCGGGGATGAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGTGAAAGGCC 2762  
Db 2700 TCGGGGATGAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGTGAAAGGCC 2759  
Qy 2763 TCGAAACACAGTGTGAAGCAGTGCAGGACACCACTGTGCCCTTGAGGACAGCATGTGGAGA 2822  
Db 2760 TCGAAACACAGTGTGAAGCAGTGCAGGACACCACTGTGCCCTTGAGGACAGCATGTGGAGA 2819  
Qy 2823 TTGCACAGCGGAGCTCTGAGTCATGTGTGAGCAACATGAGCAGTGTGTGACTC 2882  
Db 2820 TTGCACAGCGGAGCTCTGAGTCATGTGTGAGCAACATGAGCAGTGTGTGACTC 2879  
Qy 2883 CAATGCCCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGATACGATGAGCAC 2942  
Db 2880 CAATGCCCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGATACGATGAGCAC 2939  
Qy 2943 CTGCCCCCTGAAATTTGTTGAGGCTACTGTACCTGTAGTCATTTGCTTGGAGCAACGAGG 3002  
Db 2940 CTGCCCCCTGAAATTTGTTGAGGCTACTGTACCTGTAGTCATTTGCTTGGAGCAACGAGG 2999  
Qy 3003 CTGTGGCTGTGTACTGTATCCAGCAATCTGCAAGGAAATGCAATGAGAGGTTCCCTA 3062  
Db 3000 CTGTGGCTGTGTACTGTATCCAGCAATCTGCAAGGAAATGCAATGAGAGGTTCCCTA 3059  
Qy 3063 TAAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTTACAGGAAATTTCTATCCAGGCCCT 3122  
Db 3060 TAAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTTACAGGAAATTTCTATCCAGGCCCT 3119  
Qy 3123 GCTCAATTCAGCATGTGTCTAGAGGACAGCATCAATGGCTCTTCAATTCACATGTC 3182  
Db 3120 GCTCAATTCAGCATGTGTCTAGAGGACAGCATCAATGGCTCTTCAATTCACATGTC 3179  
Qy 3183 AGCTTGCAATGCAACGCGCACAGTAATGATCAATCAGAGCATCTGTGAGAGTGTGA 3242  
Db 3180 AGCTTGCAATGCAACGCGCACAGTAATGATCAATCAGAGCATCTGTGAGAGTGTGA 3239  
Qy 3243 GAACCTGACACAGGCAAGCAGTGGAGACCTGTGATATCTGGCTTCTACGGTATCCAC 3302  
Db 3240 GAACCTGACACAGGCAAGCAGTGGAGACCTGTGATATCTGGCTTCTACGGTATCCAC 3299  
Qy 3303 CAATGAGGGAATGTGAGGCATGCAATGCAATGGCAGCGCTCTGTGTGCAACACCAA 3362

Db 3300 CAATGAGGGAATGTGAGGCATGCAATGGCAGCGCTCTGTGTGCAACACCAA 3359  
Qy 3363 CACGGCAAGTGTCTCTGCACCAACCAAGGCGCTCAGGGGAGAGTGCACGCTATGTGA 3422  
Db 3360 CACGGCAAGTGTCTCTGCACCAACCAAGGCGCTCAGGGGAGAGTGCACGCTATGTGA 3419  
Qy 3423 GGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTATCTCTTAT 3482  
Db 3420 GGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTATCTCTTAT 3479  
Qy 3483 TGAATCATGATTCAGCTTCTATCCAGGAAGATGATCTCTATACAGCTATCAA 3542  
Db 3480 TGAATCATGATTCAGCTTCTATCCAGGAAGATGATCTCTATACAGCTATCAA 3539  
Qy 3543 TTTTGTGGCTACTCTCTGACGAAACAAACAGGATTTGGACATGTTTCATCAATGCCCTCAA 3602  
Db 3540 TTTTGTGGCTACTCTCTGACGAAACAAACAGGATTTGGACATGTTTCATCAATGCCCTCAA 3599  
Qy 3603 GAATTTCAACCTCAACATCACCTGGGCTGCAGTTTCTCAGCTGGAAACCCAGGCTGGAGA 3662  
Db 3600 GAATTTCAACCTCAACATCACCTGGGCTGCAGTTTCTCAGCTGGAAACCCAGGCTGGAGA 3659  
Qy 3663 AGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGA 3722  
Db 3660 AGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGA 3719  
Qy 3723 GAATTTGTGATTTTCGCAACCCAAATATCACTTTCTTTGTTTATGTGAGTAATTTTCA 3782  
Db 3720 GAATTTGTGATTTTCGCAACCCAAATATCACTTTCTTTGTTTATGTGAGTAATTTTCA 3779  
Qy 3783 CTGGCCCATCAAAATTCAGATTTGCTCTCTCAGCACAGCAATTTTATGAGACCTGATCA 3842  
Db 3780 CTGGCCCATCAAAATTCAGATTTGCTCTCTCAGCACAGCAATTTTATGAGACCTGATCA 3839  
Qy 3843 GTTCTTCGTGACTTTCTTCAGTTGTTTCTCTCTCTTTGCTCTCTGCTGCTGCTGCTG 3902  
Db 3840 GTTCTTCGTGACTTTCTTCAGTTGTTTCTCTCTCTTTGCTCTCTGCTGCTGCTGCTG 3899  
Qy 3903 GAAGATCAAAACAAAGTTGTTGGCCCTCCAGAGTGAAGCACTTTCTCGAGAGATGCA 3962  
Db 3900 GAAGATCAAAACAAAGTTGTTGGCCCTCCAGAGTGAAGCACTTTCTCGAGAGATGCA 3959  
Qy 3963 ACAGATGCGCAGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4022  
Db 3960 ACAGATGCGCAGCGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4019  
Qy 4023 TCTGATCTTATTTGGGGGAGTATTAAGACTGTTTCCAAACCCATGTCATGAGGCGGTG 4082  
Db 4020 TCTGATCTTATTTGGGGGAGTATTAAGACTGTTTCCAAACCCATGTCATGAGGCGGTG 4079  
Qy 4083 TTTTGGCAACAAAGCGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4142  
Db 4080 TTTTGGCAACAAAGCGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4139  
Qy 4143 CATCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4202  
Db 4140 CATCTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4199  
Qy 4203 CGAGATGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4262  
Db 4200 CGAGATGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4259  
Qy 4263 CCTTCACAGCCTGGGACCTGCACTGA 4290  
Db 4260 CCTTCACAGCCTGGGACCTGCACTGA 4287

## RESULT 4

AAZ52275

ID AAZ52275 standard; cDNA; 4068 BP.

XX

AC AAZ52275;

XX







Db 661 GCACATCAGGTTATGCTTGTGCTATTTTGTAGTGATGCTGCTTAATAATTCAGCTGA 720  
QY 721 TTTAATATTTACTCAGATTTTATGATGCTCAATATGCTCAATATGCTCAGGCGGAGGAGTGT 780  
Db 721 TTTAATATTTACTCAGATTTTATGATGCTCAATATGCTCAGGCGGAGGAGTGT 780  
QY 781 AAGATCAGTAATAGCAGCGAACTGTTGAATGTGAATGTTCTGAAATCTGAAAGAGGTGAA 840  
Db 781 AAGATCAGTAATAGCAGCGAACTGTTGAATGTGAATGTTCTGAAATCTGAAAGAGGTGAA 840  
QY 841 GCATGTGATCTCCCTCAGTGTACAGCAACTGTGTGTTTCTCCTCATCGAGGATCTGCAAT 900  
Db 841 GCATGTGATCTCCCTCAGTGTACAGCAACTGTGTGTTTCTCCTCATCGAGGATCTGCAAT 900  
QY 901 TCAAGTGTGTCAGAGGATGCTCCTGCTTCTCAGACTGCGAGGCTCCTGAGTGTTCAGTT 960  
Db 901 TCAAGTGTGTCAGAGGATGCTCCTGCTTCTCAGACTGCGAGGCTCCTGAGTGTTCAGTT 960  
QY 961 CTTGTACCACTCAGATTAACATGTTTGTGACTCGAGGAAATATCTAACTTAAAGCTCCCC 1020  
Db 961 CTTGTACCACTCAGATTAACATGTTTGTGACTCGAGGAAATATCTAACTTAAAGCTCCCC 1020  
QY 1021 AGAGCATCTCAATAAGCTGTGCTCAATGGAACATATATGTCGGTGTGAGGATATATG 1080  
Db 1021 AGAGCATCTCAATAAGCTGTGCTCAATGGAACATATATGTCGGTGTGAGGATATATG 1080  
QY 1081 TTCAACCACTCAGATTAACATGTTTGTAGCTATGACCTTGTCTTAGGGAGTGGCTT 1140  
Db 1081 TTCAACCACTCAGATTAACATGTTTGTAGCTATGACCTTGTCTTAGGGAGTGGCTT 1140  
QY 1141 CCACCTAAACCTTCTGTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
Db 1141 CCACCTAAACCTTCTGTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200  
QY 1201 AAGGATAAAATTTACATGATGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260  
Db 1201 AAGGATAAAATTTACATGATGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260  
QY 1261 TTGAGATTTTTCATTAATGAGTCAATGAGTGTGTTGAGTGTGTTGAGTGTGTTGAGTGT 1320  
Db 1261 TTGAGATTTTTCATTAATGAGTCAATGAGTGTGTTGAGTGTGTTGAGTGTGTTGAGTGT 1320  
QY 1321 CAGTATGAGTGTGTTGAGTCAATGAGTGTGTTGAGTGTGTTGAGTGTGTTGAGTGTGTTG 1380  
Db 1321 CAGTATGAGTGTGTTGAGTCAATGAGTGTGTTGAGTGTGTTGAGTGTGTTGAGTGTGTTG 1380  
QY 1381 ATGCTGTGCTATTTTGTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
Db 1381 ATGCTGTGCTATTTTGTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440  
QY 1441 GATTTGGATAGACACATGAGTATATTAACACCCAGGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
Db 1441 GATTTGGATAGACACATGAGTATATTAACACCCAGGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
QY 1501 TACGGCCATAGCAGTGTGTTTACGACCATAGGACCCAGGCTGCTGCTGCTGCTGCTGCTG 1560  
Db 1501 TACGGCCATAGCAGTGTGTTTACGACCATAGGACCCAGGCTGCTGCTGCTGCTGCTGCTG 1560  
QY 1561 AAGGCTTTCAGTGTGCTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
Db 1561 AAGGCTTTCAGTGTGCTATGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
QY 1621 ACCCAGATGTGACCACTTTTAAAGACAGCCGATTTTTCGCTTACTTGTGACACAGCTGTG 1680  
Db 1621 ACCCAGATGTGACCACTTTTAAAGACAGCCGATTTTTCGCTTACTTGTGACACAGCTGTG 1680  
QY 1681 ATAGTGTGGAACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
Db 1681 ATAGTGTGGAACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740  
QY 1741 CATGGCCCAATGCTTCTCTTCAATTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800

Db 1741 CATGGCCCAATGCTTCTCTTCAATTTTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800  
QY 1801 TCAGTGTGCTTCCAGAGCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860  
Db 1801 TCAGTGTGCTTCCAGAGCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860  
QY 1861 TTACACAACAGCAGCACCATGATGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920  
Db 1861 TTACACAACAGCAGCACCATGATGTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920  
QY 1921 CTGGTATTCACCTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTGTACAGCA 1980  
Db 1921 CTGGTATTCACCTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTGTACAGCA 1980  
QY 1981 GGAACCTGGTATTCGGTGTGTTGGAACACAGGCTGCTCTCAGTGTATCTCGTGGCGCTG 2040  
Db 1981 GGAACCTGGTATTCGGTGTGTTGGAACACAGGCTGCTCTCAGTGTATCTCGTGGCGCTG 2040  
QY 2041 GCACCTGATGAACAGAGAAAGTTAAATCAGATGTTTTTCCAAAAGAACTCTTTGAC 2100  
Db 2041 GCACCTGATGAACAGAGAAAGTTAAATCAGATGTTTTTCCAAAAGAACTCTTTGAC 2100  
QY 2101 CATGACAGATGTGACAGCAGACACAGATTTGTTACAGCTGTACAGCCAAACACCAATGACTGC 2160  
Db 2101 CATGACAGATGTGACAGCAGACACAGATTTGTTACAGCTGTACAGCCAAACACCAATGACTGC 2160  
QY 2161 CACTGTGCAATGACCAATTTGCTCCAGGAACACAGCTGCTCAGAGGCGCAGATCTCC 2220  
Db 2161 CACTGTGCAATGACCAATTTGCTCCAGGAACACAGCTGCTCAGAGGCGCAGATCTCC 2220  
QY 2221 ATTTTGTAGTATGAGAAATTTGCTCCAGGAATACCCCATGCTACTCTGTTAAAGAGAC 2280  
Db 2221 ATTTTGTAGTATGAGAAATTTGCTCCAGGAATACCCCATGCTACTCTGTTAAAGAGAC 2280  
QY 2281 AGCTGACAGATGTGCTCCAGCAGAACTGCTCAGAGTGGAGCCCGGAATCAGAGTGC 2340  
Db 2281 AGCTGACAGATGTGCTCCAGCAGAACTGCTCAGAGTGGAGCCCGGAATCAGAGTGC 2340  
QY 2341 ATGCTGCTGCTCCAGAAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400  
Db 2341 ATGCTGCTGCTCCAGAAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400  
QY 2401 AAAATTTACTTACTGCTCCAGGAGAAATTTATGCAATGCTAAATTTGCTGTAGGAACACCAAT 2460  
Db 2401 AAAATTTACTTACTGCTCCAGGAGAAATTTATGCAATGCTAAATTTGCTGTAGGAACACCAAT 2460  
QY 2461 GCTCTTTTGGCTTCTCTTACACCCAGAGAGAGTGTGCTTCTTAAAGCAGCTGCGA 2520  
Db 2461 GCTCTTTTGGCTTCTCTTACACCCAGAGAGAGTGTGCTTCTTAAAGCAGCTGCGA 2520  
QY 2521 ATAAATGCTGCTATCTCAGAGCATGCTCAAGCTTAAACCCCATGCTGCTGCTGCTGCTG 2580  
Db 2521 ATAAATGCTGCTATCTCAGAGCATGCTCAAGCTTAAACCCCATGCTGCTGCTGCTGCTG 2580  
QY 2581 AAGATCAATGCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640  
Db 2581 AAGATCAATGCTCTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640  
QY 2641 CAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700  
Db 2641 CAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700  
QY 2701 ACTCGGGAATGAAAGCTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
Db 2701 ACTCGGGAATGAAAGCTGCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760  
QY 2761 CTTGCAAAACACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820  
Db 2761 CTTGCAAAACACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820  
QY 2821 GATTGCAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880  
Db 2821 GATTGCAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880







2400	AAATTACTCTGCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGACCAACAATGC	2459	3543	TTTTGTGGCTACTCTGACGAAAGAAACAGGGAATTTGGACATGTTTCATCAATGCGCTCCAA	3600
2463	CTTTTGGCTTCTTTACAAACGAGAGAGTGAATTTGCTTAAAGCAGCTGCGAAT	2522	3540	TTTTGTGGCTACTCTGACGAAAGAAACAGGGAATTTGGACATGTTTCATCAATGCGCTCCAA	3599
2460	CTTTTGGCTTCTTTACAAACGAGAGAGTGAATTTGCTTAAAGCAGCTGCGAAT	2519	3603	GAATTTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGA	3662
2523	ATGCGAGTCACTCAGACGATGTCGAAGCTCACTTACCCCAAGGCTCGGCTTCGGAA	2582	3600	GAATTTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGA	3659
2520	AATGCAAGTCACTCAGACGATGTCGAAGCTCACTTACCCCAAGGCTCGGCTTCGGAA	2579	3663	AGAGATGCGCTGTTTTCACAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGA	3722
2583	GATCAATGTTGCTTCTGCTGCGGAAGATATGTCCTCCCATTTTCAAAATAGTTTACTPACA	2642	3660	AGAGATGCGCTGTTTTCACAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGA	3719
2580	GATCAATGTTGCTTCTGCTGCGGAAGATATGTCCTCCCATTTTCAAAATAGTTTACTPACA	2639	3723	GAAAGTTTCAATTTTCCCAACCAACCAATATACCTTTCTTTGTTTATGTCAAGTAATTTTCAC	3782
2643	GTGGATGCGCTGTGAGCCAGCATGCTCGGATTTCTGTGGAATTTTATCAGAACCCAGTAC	2702	3720	GAAAGTTTCAATTTTCCCAACCAACCAATATACCTTTCTTTGTTTATGTCAAGTAATTTTCAC	3779
2640	GTGGATGCGCTGTGAGCCAGCATGCTCGGATTTCTGTGGAATTTTATCAGAACCCAGTAC	2699	3783	CTGGCCCATCAAAATTCAGAT	3803
2703	TCGGGACTGAAGGCTGCAACCTGCAATCAACCCATCAATGCTAGTGTCTGTGAAGGCC	2762	3780	CTGGCCCATCAAAATTCAGAT	3800
2700	TCGGGACTGAAGGCTGCAACCTGCAATCAACCCATCAATGCTAGTGTCTGTGAAGGCC	2759	RESULT 7		
2763	TGCAAAACCAAGTCTTAAGCAGTGCAGGACACCAATGTCCTTTGAGGACAGCATGTGGAGA	2822	ADB53249		
2760	TGCAAAACCAAGTCTTAAGCAGTGCAGGACACCAATGTCCTTTGAGGACAGCATGTGGAGA	2819	ID	ADB53249 standard; DNA; 8739 BP.	
2823	TTGACACGCGGAGCTCTGAGTGCATGTGTCAGCAGCAACATGAAGCAGTGTGTGGACTC	2882	XX	ADB53249;	
2820	TTGACACGCGGAGCTCTGAGTGCATGTGTCAGCAGCAACATGAAGCAGTGTGTGGACTC	2879	AC	XX	
2883	CAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAC	2942	DT	04-DEC-2003 (first entry)	
2880	CAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAC	2939	XX	Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3791.	
2943	CTGCGCCCTGAAAATTTTCAGGCTACTGTCACCTGTAGTCAATGCTTGGAGCAACAGG	3002	DE	toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;	
2940	CTGCGCCCTGAAAATTTTCAGGCTACTGTCACCTGTAGTCAATGCTTGGAGCAACAGG	2999	KW	toxicity marker; toxicity progression; drug screening;	
3003	CTGTGGCTGTTACTGATCCAGCAATATCTGGCAAGGGAATGCATAGAGGTTTCTTA	3062	KM	primary rat hepatocyte toxicity modelling; gene; ds.	
3000	CTGTGGCTGTTACTGATCCAGCAATATCTGGCAAGGGAATGCATAGAGGTTTCTTA	3059	XX	Rattus norvegicus.	
3063	TAAAGCACAGTGAAGATGCTTTCGACGCGCTCAGGAATTTCTATCACAGCGCCT	3122	XX	WO2003065993-A2.	
3060	TAAAGCACAGTGAAGATGCTTTCGACGCGCTCAGGAATTTCTATCACAGCGCCT	3119	XX	14-AUG-2003.	
3123	GCTCAATTCAGCATGTGCTAGAGCAGCAGATACAACTGGTCTTTTCAATTCATCTGCC	3182	XX	04-FEB-2003; 2003WO-US0003482.	
3120	GCTCAATTCAGCATGTGCTAGAGCAGCAGATACAACTGGTCTTTTCAATTCATCTGCC	3179	XX	04-FEB-2002; 2002US-0353171P.	
3183	AGCTTGCCAAATGCAACGGCCACAGTAATGCAATCAATCAGACATCTGTGAGAGTGTGA	3242	XX	13-MAR-2002; 2002US-0363534P.	
3180	AGCTTGCCAAATGCAACGGCCACAGTAATGCAATCAATCAGACATCTGTGAGAGTGTGA	3239	XX	08-APR-2002; 2002US-0370248P.	
3243	GAACCTGACACAGGCAAGCACTGCGAGACCTGCATATCTGCTTCTACGGTGATCCAC	3302	XX	10-APR-2002; 2002US-0371134P.	
3240	GAACCTGACACAGGCAAGCACTGCGAGACCTGCATATCTGCTTCTACGGTGATCCAC	3299	XX	10-APR-2002; 2002US-0371135P.	
3303	CAATGAGGGAATGTCAGCCATGCAAGTGCATATGGGACCGGCTCTGTGCAACACCAA	3362	XX	10-APR-2002; 2002US-0371150P.	
3300	CAATGAGGGAATGTCAGCCATGCAAGTGCATATGGGACCGGCTCTGTGCAACACCAA	3359	XX	11-APR-2002; 2002US-0371413P.	
3363	CACGGGCAATGCTTCTGCAACCAAGGGCGTCAAGGGGACAGTGCAGCTATGTGA	3422	XX	19-APR-2002; 2002US-0373601P.	
3360	CACGGGCAATGCTTCTGCAACCAAGGGCGTCAAGGGGACAGTGCAGCTATGTGA	3419	XX	19-APR-2002; 2002US-0373602P.	
3423	GGTAGAATTCGATACCAAGGAACCTCTCAGAGCAATGTTATATCTCTCTTAT	3482	XX	19-APR-2002; 2002US-0374139P.	
3420	GGTAGAATTCGATACCAAGGAACCTCTCAGAGCAATGTTATATCTCTCTTAT	3479	XX	08-MAY-2002; 2002US-0378370P.	
3483	TGACTATCAGTTACCTTTTGTCTATCCAGGAGATGATCGCTATTACAGCTATCAA	3542	XX	09-MAY-2002; 2002US-0378652P.	
3480	TGACTATCAGTTACCTTTTGTCTATCCAGGAGATGATCGCTATTACAGCTATCAA	3539	XX	03-MAY-2002; 2002US-0378653P.	



1872 AGTTCTCCAGAGACCTGAGCTCCATCAGATGTCAACCGATTTGGCCACTCAGCAGTCTT 1931  
1863 ACACAAAGCAGCAGCATGTATGTTTCGGTGTGTTTCAATAGTCTCTCTCCTCAGCAGCATCTT 1922  
1932 GCACAAAGCAGCAGCATGTATGTTTCGGGCGGCTTCAACAGCCTCTCTCTCAGTGCAGTCTT 1991  
1923 GGTATTCACCTCGGAAACAGTGTGATGCGCATCGAGTGAAGCGCTGTTGTTAGCAGAGG 1982  
1992 AGTCTTCACTTCGGAGCAGTGTGATGACACCGCAGCGAAGCTGCTTGTGTGGCAGCAGG 2051  
1983 ACCTGGTATTCGGGTGTGTGGACACAGAGGTCTCTCAGTGTATCTCTGCGGCGCTGGC 2042  
2052 ACCTGGCATCCGGTGTCTGTGGGACACAGATGCTCTCGATGTACCTCTCTGGGAGTTGGC 2111  
2043 AACTGATGAACAAGAAAGATTAAATCAGAAATGTTTTTCCAAAAAAGAACTCTTGACCA 2102  
2112 AACTGAAGAACAGCAGAAAGTTAAATCAGATGTTTCTCTAAAGAACCCCTTGACCA 2171  
2103 TGACAGATGTGACAGCAGCAGCATGTTTACAGCTGTACAGCCACCAACCAATGACTGCCA 2162  
2172 TGACAGATGTGACAGCAGCAGCATGTTTACAGCTGTGACAGCCCAATACCAATGACTGCCA 2231  
2163 CTGGTGAATGACCAATGTGTCCCGAGAACCAAGCTGTCTCAGAGGCCAGATCTCCAT 2222  
2232 CTGGTGCATGATCACTGTGTCTGTGAACCAAGCTGTGACAGAGGCCAGATCTCCAT 2291  
2223 TTTTAGGTATGAGAAATTTGCCCAAGGATTAACCCCATGTATCTGTTAAAGAAAGACAG 2282  
2292 TGCCAAAGTATGACAAATTTGCCCAAGGATTAACCCCATGTATCTGCTCAATGAAGAAACAG 2351  
2283 CTGACAGAGCTGTCCCTGTGACCAAGACTGCGAGTGGGAGCCCGGGAATCAGGAGTGCAT 2342  
2352 CTGACAGAGCTGTCCCTGTAGTACAGAAATGCGAGTGGGAACCTCGAAATCAAGAGTGCAT 2411  
2343 TGCCCTGCCGAAATATCTGTGCAATGTGCGCAATTTGGTGGAAACTCATGTTTGA 2402  
2412 CGCCCTGCCGAAATATCTGTGCAATGTGCGCAATTTGGTGGAAACTCATGTTTGA 2471  
2403 AATTACTACTGCCAAGAGGATTTAGCAATGCTTAATTTGTTCTGTAGGAGCAGCATGTC 2462  
2472 AATCACTACTCTAAGGAAATTTAGCAATGCTTAATTTGTTCTGTAGGAGCAGCATGTC 2531  
2463 CCTTTGGCTTCTCTTACAAACCCAGAAAGGATGAGAAATTTGCTCTTAAGCAGCTGCGAAT 2522  
2532 CTTTTGGCTTCTGCTAGTCCAGAGAGAGTGTGAGTTGTTCTTAAGCAGCTTCTGAT 2591  
2523 AATGAGTCTCTCAGAGCATGTGCAAGCTCACTTAACCCCATGGGTGCGCTTCGGAA 2582  
2592 AATGCAATCATCTCAGAGCAGCTCCAAAGCTCACTCTGACTCCATGGGTGCTCTTGGAA 2651  
2583 GATCAATGTGCTTACTGTGCTGGAGATATGTCCTTCAATTTTCAAAATAGTTTACTACA 2642  
2652 GATCAATGTATCTTACTGTGCTGGAGGATATGTTCTCCATTTCAAAATAGTTTGTGCA 2711  
2643 GTGGATGCCGTCTGAGCCAGTGTGCTGAGTCTGTGGAATTTTATCAGAACCCAGTAC 2702  
2712 GTGGATGCCATCTGAGCCAGTGTGCTGCGGCTTCTGTGGGATCTTGTGAGAGCTAGTAC 2771  
2703 TCGGGAGCTGAGGCTGCAACCTGCAATCAACCCCATCAATGGTGTGCTGTGGAAGGCC 2762  
2772 TCGGGAGCTGAGGCTGCAACCTGCAATCAACCCCATCAATGGGAGTGTGTTGGAAGGCC 2831  
2763 TGCAAAACCAAGTGTGAGAGCTGCGGACACCAATGTGCTTGGAGCAGCATGTGGAGA 2822  
2832 TGCAAAACCAAGTGTGAGAGCTGCGGACACCCCTGTGCTTGGAGCAGCATGTGGAGA 2891  
2823 TTGCAACAGCGGCGAGCTCTGAGTGTGATGTGTGAGCAACATGAGCAGTGTGTGAGCTC 2882  
2892 GTGCACCAAGCAGCAGCTCCGAATGATGTGTGAGCAACATGAGCAGTGTGCTGAGCTC 2951  
2883 CAATGCCATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATCAGATGAGCAC 2942

2952 CAATGCTTACGTAGGCTCTCTTTCTTTGGCCAGTGTATGGAGTGGTATATAAATGAGCAG 3011  
2943 CTGCCCCCTCGAAAAATTTGTTTCAAGGCTACTGTACTTGTAGTCAATGCTTTGGAGCAACCCAGG 3002  
3012 CTGCCCCCTCGAAAAATTTGCTCTGGCTACTGTACTTGTACCTGCGAGCCATTTGCTTTGGAGCAGCCAGG 3071  
3003 CTGTGGCTGGTGTACTGATCCAGCAATACTTGGCAAGGGAATGATAGAGGTTTCTTA 3062  
3072 CTGTGGTGGTGTACTGATCTTAGCAATACTTGGAAAGGCAAAATGATTTAGGGAGACTTA 3131  
3063 TAAAGACCAAGTGAAGATGCTTTTCGCAAGCCCTTACAGAGAAATTTCTATCCACAGCCCTT 3122  
3132 TAAAGACCCGTGAAGATGCTTTTCAATGCTCTTACAGGAATGTACCCACAGCCCTT 3191  
3123 GCTCAATTTCCAGCATGTGTCTAGAGCAGCAGATGATCAACTGGTCTTTTCAATTCAGTGTCC 3182  
3192 TCTGAATCTCCAGCATGTGTCTAGAGCAGCAGATGATCAACTGGTCTTTTATTTCACTGTCC 3251  
3183 AGCTTGCCCAATCAAGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAAGTGA 3242  
3252 AGCTTGCCAGTGCATGAGCAGCAGATGATCAACTGGTCTTTTATTTCACTGTCC 3311  
3243 GAACCTGACCCAGCAGCAAGCACTGCGAGACCTGTCATATCTGGCTTCTTACGGTGTATCCAC 3302  
3312 CGACCTGACCCAGCAGCAAGCACTGCGAGACCTGTCATATCTGGCTTCTTATGGTGAOCCGAC 3371  
3303 CAATGGAGGAATGTCAAGCATGCAAGTGCATGCGGCAAGCTCTCTGTGCAACACCA 3362  
3372 TAATGAGGCAATGTCAAGCATGCAAGTGCATGCGGCAAGCTCTCTGTGCAACACCA 3431  
3363 CACGGCAAGTGTCTTGTGCAACCAAGGCGCTCAAGGGGAGCAGTGTGCCATGTGA 3422  
3432 CACTGGCAAGTGTCTTGTGACCAAGGCGCTCAAGGGAGAGGAGTGTGCCATGTGA 3491  
3423 GGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATTTATCTTCTTAT 3482  
3492 GGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATTTATCTTCTTAT 3551  
3483 TGACTATCAGTTTCACTTTTAGTCTATCCAGGAAGATGATCCGCTATTACACAGCTATCAA 3542  
3552 TGACTATCAGTTTCACTTTTAGTCTATCCAGGAAGATGATCCGCTATTACACAGCTATCAA 3611  
3543 TTTTGTGGCTACTCTGAGCAACAAACAGGAGTTTGGACATGTTTCATCAATGCTCCAA 3602  
3612 TTTTGTGGCTACTCTGAGCAACAAACAGGAGTTTGGACATGTTTCATCAATGCTCCAA 3671  
3603 GAATTTCACTTCAACATCAGCTGGCTGCGAGTTTCTCAGCTGGAAACCCAGCTGGAGA 3662  
3672 AACTTTCAACCTCAACATCAGCTGGCTGCGAGTTTCTCAGCTGGAAACCCAGCTGGAGA 3731  
3663 AGAGATGCTGTGTTTCAAAAAACCAATTAAGGAGTACAAAGATAGTTTCTCTAATGA 3722  
3732 AGAGGTGCTGTGTTTCAAAAAACCAATTAAGGAGTACAAAGATAGTTTCTCTAATGA 3791  
3723 GAAGTTTGAATTTTGGCAACCAACCAATATCACTTTCTTTGTTATGTAGTAAATTTAC 3782  
3792 GAATTTGAATTTTGGCAACCAACCAATATCACTTTCTTTGTTATGTAGTAAATTTAC 3851  
3783 CTGGCCCATCAAAATTCAGATGCTCTCTCAGACAGCAATTTTATCGAAGCTGTGACA 3842  
3852 TTGGCCCATCAAAATTCAGATGCTCTCTCAGACAGCAATTTTATGAGCTGTGACA 3911  
3843 GTTCTTCTGTGACTTCTTCAAGTGTGTTTCTCTTTGCTCTCTGCTGGCTGTGTTG 3902  
3912 GTTCTTCTGTGACTTCTTCAAGTGTGTTTCTCTCTCTGCTGCTGGGCGAGTGTGCTG 3971  
3903 GAAGATCAACAAAGTGTGTTGGCCCTCCAGAGTGTAGAGAGCACTTCTTTCAGAGATGCA 3962  
3972 GAAGATCAACAGAGTGTGTTGGCATCCAGAGGAGAGAGCACTTCTTTCGGGAGTGA 4031  
3963 ACAGATGGCCAGCGCTCTTGGCTCTCTGTAATGTGCGCTTGGAAACAGATGAGAGCC 4022  
4032 ACAGATGGCCAGCGCGCTCTTGGCTCTCTGTAATGTGCGCTTGGAAACAGATGAGAGCC 4091



Db 948 TGTAGATATGGTCATCTTTGGCAATTATACAGGATAAAATTTACATGATATGGAGGAA 1007  
Qy 1230 AATTGATCAAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCATTTCAATTAATGAGTC 1289  
Db 1008 AATTGATCAAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCATTTCAATTAATGAGTC 1067  
Qy 1290 ATGGGTGTTGTTGACCCCTTAAGCAAGAGAGAGATATGAGTGGTGGGCACTCTGCACA 1349  
Db 1068 ATGGGTGTTGTTGACCCCTTAAGCAAGAGAGAGATATGAGTGGTGGGCACTCTGCACA 1127  
Qy 1350 CATTTGTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACCTGCCCTCT 1409  
Db 1128 CATTTGTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACCTGCCCTCT 1187  
Qy 1410 CTATGGATATATAAGCAATGTGAGGAATATGATTTGGTAAGAACACACATGGAGTATATT 1469  
Db 1188 CTATGGATATATAAGCAATGTGAGGAATATGATTTGGTAAGAACACACATGGAGTATATT 1247  
Qy 1470 ACACACCCAGGGTCCCTTTGTGCAAGGGGTACGGCCATAGCAGTGGTTTACGACCATAG 1529  
Db 1248 ACACACCCAGGGTCCCTTTGTGCAAGGGGTACGGCCATAGCAGTGGTTTACGACCATAG 1307  
Qy 1530 GACACGGCCCTATACGTTTCATGCTGGCTACAAAGGCTTTTCAGTGCCCAATAAGTACCGCT 1589  
Db 1308 GACACGGCCCTATACGTTTCATGCTGGCTACAAAGGCTTTTCAGTGCCCAATAAGTACCGCT 1367  
Qy 1590 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGACCAATTTCTTAAGGACAG 1649  
Db 1368 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGACCAATTTCTTAAGGACAG 1427  
Qy 1650 CCGATTTTCCGTTTACTTGCACACAGCTGTGATGAGTGGGAACCAATGCTGTGTGTTGG 1709  
Db 1428 CCGATTTTCCGTTTACTTGCACACAGCTGTGATGAGTGGGAACCAATGCTGTGTGTTGG 1487  
Qy 1710 GGGAAACACACAAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTTTCAATTT 1769  
Db 1488 GGGAAACACACAAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTTTCAATTT 1547  
Qy 1770 CATGGCCCTATGACATTTGCGCTGTGACCGCTGGTCAGTCTTCCAGACCTGATCTCCACCA 1829  
Db 1548 CATGGCCCTATGACATTTGCGCTGTGACCGCTGGTCAGTCTTCCAGACCTGATCTCCACCA 1606  
Qy 1830 TGATGTCAACAGATTTGGCCATT-CAGCAGTCTTTACACACAGCACCATGATGTGTGTTGG 1888  
Db 1607 TGATGTCAACAGATTTGGCCATTCCAGCAGTCTTTACACACAGCACCATGATGTGTGTTGG 1666  
Qy 1889 GTGGTTTCAATAGTCTCTCTCCTCAGCAACATCTCTGATTTCACTCGGAAACAGTGTGATG 1948  
Db 1667 GTGGTTTCAATAGTCTCTCTCCTCAGCAACATCTCTGATTTCACTCGGAAACAGTGTGATG 1726  
Qy 1949 CGCATCGAGTGAAGCCGCTGTTTACAGCAGGACCTGGTATTTCCGTGTGTGTGGAACA 2008  
Db 1727 CGCATCGAGTGAAGCCGCTGTTTACAGCAGGACCTGGTATTTCCGTGTGTGTGGAACA 1786  
Qy 2009 CAGGGTCTGTCTCAGTGTATCTCGTGGCGCTGGCACTGATGAACAAAGAAAGTTAA 2068  
Db 1787 CAGGGTCTGTCTCAGTGTATCTCGTGGCGCTGGCACTGATGAACAAAGAAAGTTAA 1846  
Qy 2069 AATCAGATGTTTTTCAAGAAGACTCTTGACCATGACAGATGTGACCCACACAGATT 2128  
Db 1847 AATCAGATGTTTTTCAAGAAGACTCTTGACCATGACAGATGTGACCCACACAGATT 1906  
Qy 2129 GTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTGCAATGACCAATTTGTCTCCCA 2188  
Db 1907 GTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTGCAATGACCAATTTGTCTCCCA 1966  
Qy 2189 GGAACCAAGCTGTCTAGAGGCCAGATCTCCATTTTATGATGAGAAATTTGCCCAAGG 2248  
Db 1967 GGAACCAAGCTGTCTAGAGGCCAGATCTCCATTTTATGATGAGAAATTTGCCCAAGG 2026  
Qy 2249 ATAAACCCCTGCTACTTGTAAACAGAGAACCCAGCTGCAGGAGTGTGCCCTGGACAGA 2308  
Db 2027 ATAAACCCCTGCTACTTGTAAACAGAGAACCCAGCTGCAGGAGTGTGCCCTGGACAGA 2086

Qy 2309 ACTGCCAGTGGAGACCCCGGAATCAGGAGTGCATTTGCCCTGCCGAAATATCTGTGGCA 2368  
Db 2087 ACTGCCAGTGGAGACCCCGGAATCAGGAGTGCATTTGCCCTGCCGAAATATCTGTGGCA 2146  
Qy 2369 TTGCTCGCATTTGTTGGTGGAACTCATGTTTGAATAATTTACTACTGCGCAAGAGAAATATG 2428  
Db 2147 TTGCTCGCATTTGTTGGTGGAACTCATGTTTGAATAATTTACTACTGCGCAAGAGAAATATG 2206  
Qy 2429 ACAATGCTTAAATTTGTTCTGTAGGAACCAATATGCTTCTTTGGCTTCTTTTCAACCCAGA 2488  
Db 2207 ACAATGCTTAAATTTGTTCTGTAGGAACCAATATGCTTCTTTGGCTTCTTTTCAACCCAGA 2266  
Qy 2489 AGAGGTAGAAATTTGTTCTTAAAGCAGCTGCGAATAATGAGTGCATCTCAGAGCATGTCCA 2548  
Db 2267 AGAGGTAGAAATTTGTTCTTAAAGCAGCTGCGAATAATGAGTGCATCTCAGAGCATGTCCA 2326  
Qy 2549 AGCTCACCTTAAACCCATGCGGCTTTCGGAAGATCAATGTGTCTTACTTGTGCTGGG 2608  
Db 2327 AGCTCACCTTAAACCCATGCGGCTTTCGGAAGATCAATGTGTCTTACTTGTGCTGGG 2386  
Qy 2609 AAGATATGCTCCCAATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCAGTGTG 2668  
Db 2387 AAGATATGCTCCCAATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCAGTGTG 2446  
Qy 2669 CTGGATTTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACTGCA 2728  
Db 2447 CTGGATTTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACTGCA 2506  
Qy 2729 TCAACCCACTCAATGGTGTGTTCTGTGAAAGCCCTGCAACACACAGTGTAAAGCAGTGCC 2788  
Db 2507 TCAACCCACTCAATGGTGTGTTCTGTGAAAGCCCTGCAACACACAGTGTAAAGCAGTGCC 2566  
Qy 2789 GGACACCATGTCCTTGGAGCAGCATGTGGAGATGACACAGCGGAGCTCTGAGTGTCA 2848  
Db 2567 GGACACCATGTCCTTGGAGCAGCATGTGGAGATGACACAGCGGAGCTCTGAGTGTCA 2626  
Qy 2849 TGTGTGTGACGAAACATGAAGCAGTGTGTGGACTCCAATGCTTGTGCTTCTTCCCTT 2908  
Db 2627 TGTGTGTGACGAAACATGAAGCAGTGTGTGGACTCCAATGCTTGTGCTTCTTCCCTT 2686  
Qy 2909 TTGGCCAGTGTATGAAATGTTATGATGAGACACTGCGCCCTCTGAAATTTGTTGAGCT 2968  
Db 2687 TTGGCCAGTGTATGAAATGTTATGATGAGACACTGCGCCCTCTGAAATTTGTTGAGCT 2746  
Qy 2969 ACTGTACCTGTAGTCAATGTTGGAGCAACAGGCTGTGGTGTGTTACTGATCCACGCA 3028  
Db 2747 ACTGTACCTGTAGTCAATGTTGGAGCAACAGGCTGTGGTGTGTTACTGATCCACGCA 2806  
Qy 3029 ATACTGGCAAAAGGAAATGCAATAGAGGTTTCTATAAAGGACCAAGTGAAGATGCTTCCG 3088  
Db 2807 ATACTGGCAAAAGGAAATGCAATAGAGGTTTCTATAAAGGACCAAGTGAAGATGCTTCCG 2866  
Qy 3089 AAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATTCACAGCATGTGTAGAGG 3148  
Db 2867 AAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATTCACAGCATGTGTAGAGG 2926  
Qy 3149 ACAGCAGATACAACTGGTCTTTTCAATTCATCTGTCAGCTTGGCAATGCAACGCCACAGTA 3208  
Db 2927 ACAGCAGATACAACTGGTCTTTTCAATTCATCTGTCAGCTTGGCAATGCAACGCCACAGTA 2986  
Qy 3209 AATGATCAATCAGAGCATCTGTGAAAGTGTGAGAACTTGACCAACAGGCAAGCACTGCG 3268  
Db 2987 AATGATCAATCAGAGCATCTGTGAAAGTGTGAGAACTTGACCAACAGGCAAGCACTGCG 3046  
Qy 3269 AGACCTTGCATATCTGGCTTCTACGGTGTATCCCAATGGAGGAAATGTCAGCCATGCA 3328  
Db 3047 AGACCTTGCATATCTGGCTTCTACGGTGTATCCCAATGGAGGAAATGTCAGCCATGCA 3106  
Qy 3329 AGTGAATGGGCAACGCTCTGTGCAACACAAACAGGGCAAGTGTCTTGTGCACACCA 3388  
Db 3107 AGTGAATGGGCAACGCTCTGTGCAACACAAACAGGGCAAGTGTCTTGTGCACACCA 3166



QY 3389 AGGGCTCAAGGGGACAGTGCACGCTATGTGAGTAGAATAATCGATACCAAGAAACC 3448  
 DB |||||  
 QY 3167 AGGGCGTCAAGGGGACAGTGCACGCTATGTGAGTAGAATAATCGATACCAAGAAACC 3226  
 DB |||||  
 QY 3449 CTCTCAGAGAACATGTTATTATATCTCTTCTTATTGACTATCACTTACCTTTAGTCTAT 3508  
 DB |||||  
 QY 3227 CTCTCAGAGAACATGTTATTATATCTCTTCTTATTGACTATCACTTACCTTTAGTCTAT 3286  
 DB |||||  
 QY 3509 CCCAGGAAGATGCTGCTATTACACAGCTATCAATTTTGGGCTACTCTCGACGACAAA 3568  
 DB |||||  
 QY 3287 CCCAGGAAGATGCTGCTATTACACAGCTATCAATTTTGGGCTACTCTCGACGACAAA 3346  
 DB |||||  
 QY 3569 ACAGGGATTGGACATGTTCAATCAATGCTCCCAAGAAATTTCAACCTCAACATCACCTGGG 3628  
 DB |||||  
 QY 3347 ACAGGGATTGGACATGTTCAATCAATGCTCCCAAGAAATTTCAACCTCAACATCACCTGGG 3406  
 DB |||||  
 QY 3629 CTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCCCTGTTGTTTCAAAAACCA 3688  
 DB |||||  
 QY 3407 CTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCCCTGTTGTTTCAAAAACCA 3466  
 DB |||||  
 QY 3689 ACATTAGGCTACAAAGATAGTTTCTCTAATGAGAGATTTGATTTTCGCAACCCCAA 3748  
 DB |||||  
 QY 3467 ACATTAGGCTACAAAGATAGTTTCTCTAATGAGAGATTTGATTTTCGCAACCCCAA 3526  
 DB |||||  
 QY 3749 ATATCACTTTCTTTGTTTATGCTAGTAATTTTCACTGGCCCATCAAAATTCAGAT 3803  
 DB |||||  
 QY 3527 ATATCACTTTCTTTGTTTATGCTAGTAATTTTCACTGGCCCATCAAAATTCAGAT 3581  
 DB |||||

## RESULT 9

AAS72657  
 ID AAS72657 standard; cDNA; 3597 BP.

XX AC AAS72657;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #8461.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG08470.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
 diagnostics, forensics, gene mapping, identification of mutations  
 responsible for genetic disorders or other traits and to assess  
 biodiversity.

XX PS Claim 1; SEQ ID NO 8461; 103bp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 sequences. (I) is useful as hybridisation probes, polymerase chain  
 reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal

CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX  
 SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 80.9%; Score 3469.4; DB 5; Length 3597;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 3507; Conservative 0; Mismatches 26; Indels 2; Gaps 2;

QY 270 GTGCGGCTCAGCGGAGCGGAGGCAAGAAATGTGACCGGCCCTGTGTCAACGGCGTGC 329

DB 48 GCGCGGCAAGCGGAGCGGCTCGCGGGCAGGAGCGGGCGGCACCTGTGTCAACGGCGTGC 107

QY 330 CTGCAACCTGSCACCGCCAGTGGCTGCTGCCCGCGGCTGGTGGCGGAGCAATGCCA 389

DB 108 CTGCAACCTGSCACCGCCAGTGGCTGCTGCCCGCGGCTGGTGGCGGAGCAATGCCA 167

QY 390 GCACCTGCGGGGCGCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGCCCTGG 449

DB 168 GCACCTGCGGGGCGCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGCCCTGG 227

QY 450 AAATTTATAATACAAACGAAGTCAGCTGCTCATTTGAAGGACAGCAATAGAAAT 509

DB 228 AAATTTATAATACAAACGAAGTCAGCTGCTCATTTGAAGGACAGCAATAGAAAT 287

QY 510 GAGACTTCGTTTCAATCAATTTTGTACAGAGTAGTTGGGACCAATTTATATGTTTATGA 569

DB 288 GAGACTTCGTTTCAATCAATTTTGTACAGAGTAGTTGGGACCAATTTATATGTTTATGA 347

QY 570 TGGGACCTCAATTTATGCAACCGCTAGTTGCTGCAATTTAGTGGCCCTCATTTGCTTCAGAG 629

DB 348 TGGGACCTCAATTTATGCAACCGCTAGTTGCTGCAATTTAGTGGCCCTCATTTGCTTCAGAG 407

QY 630 AGATGGCAATGAGACTGTCCTGAGGTTGTTGCCACATCAGTTATGCTTGTGCTGCAATTT 689

DB 408 AGATGGCAATGAGACTGTCCTGAGGTTGTTGCCACATCAGTTATGCTTGTGCTGCAATTT 467

QY 690 TTTTAGTCAATGCTGCTTATAATTTTCACTGGATTTAAATTTACTTTACAGTTTGTATGTTG 749

DB 468 TTTTAGTCAATGCTGCTTATAATTTTCACTGGATTTAAATTTACTTTACAGTTTGTATGTTG 527

QY 750 TCCAAATAACTGCTCAGGCCGAGGAGTGTAAGATCAGTAATAGACGCGAACTGTGA 809

DB 528 TCCAAATAACTGCTCAGGCCGAGGAGTGTAAGATCAGTAATAGACGCGAACTGTGA 587

QY 810 ATGTGAATGTTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869

DB 588 ATGTGAATGTTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647

QY 870 CTGTGGTTTCTCTCATCGAGGCATCTGCAATTTCAAGTGTGTGACAGATGCTCTCTT 929

DB 648 CTGTGGTTTCTCTCATCGAGGCATCTGCAATTTCAAGTGTGTGACAGATGCTCTCTT 707

QY 930 CTCAGACTGGCAGGCTCCTGGATGTTTCACTTCTCTGTACACAGCTAACCCAGTCATTTTGGAC 989

DB 708 CTCAGACTGGCAGGCTCCTGGATGTTTCACTTCTCTGTACACAGCTAACCCAGTCATTTTGGAC 767

QY 990 TCGAGAGGAATATTCTTAACCTTAAAGCTCCCGCAGAGCATCTCATTAAGCTGTGTTCAATGG 1049

DB 768 TCGAGAGGAATATTCTTAACCTTAAAGCTCCCGCAGAGCATCTCATTAAGCTGTGTTCAATGG 827

QY 1050 AAACATTTATGTTGGAGATATATGTTCAACCACTCAGATTATAAACAATGTTCT 1109  
Db 828 AAACATTTATGTTGGAGATATATGTTCAACCACTCAGATTATAAACAATGTTCT 887  
QY 1110 AGCGATATGCTGCTCTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGT 1169  
Db 888 AGCGATATGCTGCTCTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGT 947  
QY 1170 TGTAGATATGTCATTTGTCGATTTATACAGGATTAATAATTTACATGATGAGGAAA 1229  
Db 948 TGTAGATATGTCATTTGTCGATTTATACAGGATTAATAATTTACATGATGAGGAAA 1007  
QY 1230 AATTGATCCAACTGGGAATGTGACCAATGATGTTGAGAGTTTTTTCACATTCATAATGATC 1289  
Db 1008 AATTGATCCAACTGGGAATGTGACCAATGATGTTGAGAGTTTTTTCACATTCATAATGATC 1067  
QY 1290 ATGGGTGTTGTTGACCCCTTAAGCAAGAGGAGTATGCACTGGTGGGCACTCTGCACA 1349  
Db 1068 ATGGGTGTTGTTGACCCCTTAAGCAAGAGGAGTATGCACTGGTGGGCACTCTGCACA 1127  
QY 1350 CATTTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGGCCCTCT 1409  
Db 1128 CATTTGTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTTGGTCACTGGCCCTCT 1187  
QY 1410 CTATGATATATAGCAATGTGACGGAATATGATTTGGATAGAAACAACATGGAGTATAT 1469  
Db 1188 CTATGATATATAGCAATGTGACGGAATATGATTTGGATAGAAACAACATGGAGTATAT 1247  
QY 1470 ACACACCCAGGTCCTGTCGAAGGGGTTACGGCCATAGCAGTGTTTTACCAACCATAG 1529  
Db 1248 ACACACCCAGGTCCTGTCGAAGGGGTTACGGCCATAGCAGTGTTTTACCAACCATAG 1307  
QY 1530 GACCAGGCGCCCTAPACGTTTCATGTTGCTCAAGGGCTTTTCACTGCCAATAGTACCGGCT 1589  
Db 1308 GACCAGGCGCCCTAPACGTTTCATGTTGCTCAAGGGCTTTTCACTGCCAATAGTACCGGCT 1367  
QY 1590 TGCAGATGATCTTACCGATATGATGTGGATACCCAGATGAGCAATCTTAAAGCACAG 1649  
Db 1368 TGCAGATGATCTTACCGATATGATGTGGATACCCAGATGAGCAATCTTAAAGCACAG 1427  
QY 1650 CCGATTTTTCGTTTACCTGACACAGCTGTGATAGTCAAGTGGAACTGCTGGTGTGG 1709  
Db 1428 CCGATTTTTCGTTTACCTGACACAGCTGTGATAGTCAAGTGGAACTGCTGGTGTGG 1487  
QY 1710 GGGAAAACACACAATGACACATCTATGAGCCATGGCGCCAAATGTTCTCTTCAGATTT 1769  
Db 1488 GGGAAAACACACAATGACACATCTATGAGCCATGGCGCCAAATGTTCTCTTCAGATTT 1547  
QY 1770 CATGSCCTATGACATTTGCTGTGACCGTGTGATGCTTCCAGACCTGATCTCCACCA 1829  
Db 1548 CATGSCCTATGACATTTGCTGTGACCGTGTGATGCTTCCAGACCTGAT-TCACCA 1606  
QY 1830 TGATGTCAACAGATTTGGCCATT-CAGCAGTCTTTACACAACAGCACCATGATGTGTGTCG 1888  
Db 1607 TGATGTCAACAGATTTGGCCATTCCAGCAGTCTTTACACAACAGCACCATGATGTGTGTCG 1666  
QY 1889 GTGGTTTCAATAGTCTCTCTCAGCGACATCTGTTATTCACCTCGGAACAGTGTGATG 1948  
Db 1667 GTGGTTTCAATAGTCTCTCTCAGCGACATCTGTTATTCACCTCGGAACAGTGTGATG 1726  
QY 1949 CGCATCGGAGTGAAGCCGCTTTGTTAGCAGCAGACCTGTTATTCGGTGTGTGTGGAACA 2008  
Db 1727 CGCATCGGAGTGAAGCCGCTTTGTTAGCAGCAGACCTGTTATTCGGTGTGTGTGGAACA 1786  
QY 2009 CAGGTCGTCTCAGTGTATCTCGTGGGGCTGGCACTGATGAACAAGAGAAAGTTAA 2068  
Db 1787 CAGGTCGTCTCAGTGTATCTCGTGGGGCTGGCACTGATGAACAAGAGAAAGTTAA 1846  
QY 2069 AATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACAGCACACAGATT 2128  
Db 1847 AATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACAGCACACAGATT 1906

QY 2129 GTTACAGCTGTACAGCCAAACCAATGATGCCACTGCTGCTGCAATGACCAATGTTGTCCTCA 2188  
Db 1907 GTTACAGCTGTACAGCCAAACCAATGATGCCACTGCTGCTGCAATGACCAATGTTGTCCTCA 1966  
QY 2189 GGAACCAACAGCTGTCTCAGAGGCCAGATCTCCATTTTATAGTATGAGAAATGCCCCAAGG 2246  
Db 1967 GGAACCAACAGCTGTCTCAGAGGCCAGATCTCCATTTTATAGTATGAGAAATGCCCCAAGG 2026  
QY 2249 ATAAACCCCATGTACTACTGTAAACAGAACAGCAGCTGACAGAGCTGTGCCCTGGAACACAGA 2308  
Db 2027 ATAAACCCCATGTACTACTGTAAACAGAACAGCAGCTGACAGAGCTGTGCCCTGGAACACAGA 2086  
QY 2309 ACTGCGAGTGGAGGCCCGGAATCAGGAGTGCATTCCTGCCGAAATATCTGTGCA 2368  
Db 2087 ACTGCGAGTGGAGGCCCGGAATCAGGAGTGCATTCCTGCCGAAATATCTGTGCA 2146  
QY 2369 TTGGCTGTCATTTGCTTGGAAACTCATGTTTGAATAATTTACTACTGCAAGAGAAATTTATG 2428  
Db 2147 TTGGCTGTCATTTGCTTGGAAACTCATGTTTGAATAATTTACTACTGCAAGAGAAATTTATG 2206  
QY 2429 ACAATGCTAAATTTGCTTGGAAACTCATGTTTGAATAATTTACTACTGCAAGAGAAATTTATG 2488  
Db 2207 ACAATGCTAAATTTGCTTGGAAACTCATGTTTGAATAATTTACTACTGCAAGAGAAATTTATG 2266  
QY 2489 AGAAGGTAGAAATTTGCTTGAAGAGCTGCGAAATTAATGAGTCAATCTCAGAGCATGTCCA 2548  
Db 2267 AGAAGGTAGAAATTTGCTTGAAGAGCTGCGAAATTAATGAGTCAATCTCAGAGCATGTCCA 2326  
QY 2549 AGCTCACTTAAACCCCATGCTGCGCTTTCGGAAGATCAATGTTGCTTCTTACTGCTGGG 2608  
Db 2327 AGCTCACTTAAACCCCATGCTGCGCTTTCGGAAGATCAATGTTGCTTCTTACTGCTGGG 2386  
QY 2609 AAGATATGCTCCCATTTACAAATAGTCTTACTACAGTGCCTGCTGAGCCAGTGTG 2668  
Db 2387 AAGATATGCTCCCATTTACAAATAGTCTTACTACAGTGCCTGCTGAGCCAGTGTG 2446  
QY 2669 CTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACTGCA 2728  
Db 2447 CTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACTGCA 2506  
QY 2729 TCAACCCCATCAATGGTGTGCTGTGAAGSCCTGCAAAACAACAGTCTAAGCAGTGGC 2788  
Db 2507 TCAACCCCATCAATGGTGTGCTGTGAAGSCCTGCAAAACAACAGTCTAAGCAGTGGC 2566  
QY 2789 GGACACCATGCTCCCTTTGAGGACAGCATGTGGAGATTTGACACAGCGGAGCTCTGAGTGA 2848  
Db 2567 GGACACCATGCTCCCTTTGAGGACAGCATGTGGAGATTTGACACAGCGGAGCTCTGAGTGA 2626  
QY 2849 TGTGCTGACGCAACATGAAGCAGTGTGGAATCCAAATGCTTATGTGGCTCTCTTCCCTT 2908  
Db 2627 TGTGCTGACGCAACATGAAGCAGTGTGGAATCCAAATGCTTATGTGGCTCTCTTCCCTT 2686  
QY 2909 TTGGCCAGTGTATGGATGATGATGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2968  
Db 2687 TTGGCCAGTGTATGGATGATGATGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2746  
QY 2969 ACTGTACCTGTAGTCAATTTCTTTGGAGCAACAGGCTGTGGTGTGCTGCTGCTGCTGCTGCTGCT 3028  
Db 2747 ACTGTACCTGTAGTCAATTTCTTTGGAGCAACAGGCTGTGGCTGTGGTGTGCTGCTGCTGCTGCT 2806  
QY 3029 ATACTGGCAAGGGAATGCAATAGAGGTTCTTAATAAGGACAGTGAAGTGCCTTCCG 3088  
Db 2807 ATACTGGCAAGGGAATGCAATAGAGGTTCTTAATAAGGACAGTGAAGTGCCTTCCG 2866  
QY 3089 AAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGG 3148  
Db 2867 AAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGG 2926  
QY 3149 ACAGCAGATACAACTGGTCTTTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3208  
Db 2927 ACAGCAGATACAACTGGTCTTTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2986  
QY 3209 AATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCAACAGCAAGCACTGGC 3268

DB 2987 AATGATCAATCAGAGCATCTGTGAGAGTGTGAGACCTGACACAGGACAGCATGCG 3046  
 QY 3269 AGACCTGATATCTGGCTTCTACCGGTGATCCCAACATGGAGGAAATGTGACGCATGCA 3328  
 DB 3047 AGACCTGATATCTGGCTTCTACCGGTGATCCCAACATGGAGGAAATGTGACGCATGCA 3106  
 QY 3329 AGTGAATGGGACGCTCTGTGACACACACGCGCAAGTGTCTGACACACCA 3388  
 DB 3107 AGTGAATGGGACGCTCTGTGACACACACGCGCAAGTGTCTGACACACCA 3166  
 QY 3389 AGGGCGTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAC 3448  
 DB 3167 AGGGCGTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAC 3226  
 QY 3449 CTCTCAGAGGACATGTTATTATATCTCTTATGACTATCAGTTACCTTTAGTCTAT 3508  
 DB 3227 CTCTCAGAGGACATGTTATTATATCTCTTATGACTATCAGTTACCTTTAGTCTAT 3286  
 QY 3509 CCCAGGAGATGATCGCTATTATACACAGCTATCAATTTTGTGGTACTCTCGACGACAAA 3568  
 DB 3287 CCCAGGAGATGATCGCTATTATACACAGCTATCAATTTTGTGGTACTCTCGACGACAAA 3346  
 QY 3569 ACAGGATTTGGACATGTTATCAATGCTCCAAAGATTTCAACCTCAACATCACCTGGG 3628  
 DB 3347 ACAGGATTTGGACATGTTATCAATGCTCCAAAGATTTCAACCTCAACATCACCTGGG 3406  
 QY 3629 CTGCCAGTTTCTCAGCTGGAACCCAGCTGGAGAGAGATGCTGTGTTTCAAAACCA 3688  
 DB 3407 CTGCCAGTTTCTCAGCTGGAACCCAGCTGGAGAGAGATGCTGTGTTTCAAAACCA 3466  
 QY 3689 ACATTAAGAGTACAAAGATAGTTTCTTAATGAGAGTTTGATTTTGGCAACCAACCCAA 3748  
 DB 3467 ACATTAAGAGTACAAAGATAGTTTCTTAATGAGAGTTTGATTTTGGCAACCAACCCAA 3526  
 QY 3749 ATATCACTTTCTTTGTTTATGTCAGTAATTTACCTGCGCCCATCAAAATTCAGAT 3803  
 DB 3527 ATATCACTTTCTTTGTTTATGTCAGTAATTTACCTGCGCCCATCAAAATTCAGAT 3581

RESULT 10  
 ADD70997  
 ID ADD70997 standard; DNA; 3597 BP.  
 XX  
 AC ADD70997;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Human attractin gene SEQ ID NO:1.  
 XX  
 KW liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;  
 .KW cytostatic; gene therapy; human; gene; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003061564-A2.  
 XX  
 PD 31-JUL-2003.  
 XX  
 PF 20-DEC-2002; 2002WO-US040718.  
 XX  
 PR 21-DEC-2001; 2001US-0341815P.  
 PR 31-DEC-2001; 2001US-0343185P.  
 XX  
 PA (GENE-) GENE LOGIC INC.  
 PA (LGBI-) LG BIOMEDICAL INST.  
 XX  
 PI Koh SS, Liu Q, Chung H, Zeng W, Lee B, Yeramilli S, Song SY;  
 XX  
 DR WPI; 2003-663343/62.  
 XX  
 PT Diagnosing liver cancer cells, useful for treating liver cancer  
 PT associated with chronic hepatitis or cirrhosis comprises detecting the

PT Level of expression in a tissue sample of one or more genes associated  
 PT with cancerous liver tissues.

Claim 1; SEQ ID NO 1; 176pp; English.

XX The present invention describes a method for diagnosing liver cancer  
 XX cells comprising detecting the level of expression in a tissue sample of  
 CC one or more genes given in the specification (see ADD70997 to ADD71105),  
 CC where differential expression of the genes is indicative of liver cancer.  
 CC Also described: (1) detecting the progression of liver cancer in a  
 CC patient; (2) monitoring the treatment of a patient with liver cancer; (3)  
 CC treating a patient with liver cancer; (4) typing a liver disease in a  
 CC patient; (5) detecting the presence or progression of liver cancer in a  
 CC patient with chronic hepatitis or cirrhosis; (6) differentiating liver  
 CC cancer related to chronic hepatitis from liver cancer related to  
 CC cirrhosis; (7) screening for an agent capable of modulating the onset or  
 CC progression of liver cancer; (8) a composition comprising at least two  
 CC oligonucleotides comprising a sequence that specifically hybridizes to  
 CC any of the genes; (9) a solid support comprising the at least two  
 CC oligonucleotides; (10) a computer system comprising a database containing  
 CC information identifying the level in liver tissue of a set of genes; (11)  
 CC a method for using the computer system to present information identifying  
 CC the expression level in tissue or cell of any of the genes; and (12) a  
 CC therapeutic agent for slowing or halting the progression of liver cancer.  
 CC The methods are useful for treating liver cancer associated with chronic  
 CC hepatitis or cirrhosis. The present sequence represents a specifically  
 CC claimed human gene sequence which is used in the exemplification of the  
 CC present invention.

XX Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 80.9%; Score 3469.4; DB 9; Length 3597;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 3507; Conservative 0; Mismatches 26; Indels 2; Gaps 2;

QY 270 GTCCGGCTCAGCCGCGAGGCGGCAAGATGTGACCGGCCCTGTGTCAACGGCGGTG 329  
 DB 48 GCGCGCGAGCGGCGGCTCGCGGGGAGGAGCGGGCGGCGGCGGCTGTGTCAACGGCGGTG 107  
 QY 330 CTGCAACCCCTGGCACCGGCGAGTGGCTGTGCCCGCGCGCTGGGTGGCGGAGCAATGCCA 389  
 DB 108 CTGCAACCCCTGGCACCGGCGAGTGGCTGTGCCCGCGCGCTGGGTGGCGGAGCAATGCCA 167  
 QY 390 GCATCGCGGGGCGCCTTCAGACTAACTGGATCTCTCGGGTTTGTGACAGATGACCTGG 449  
 DB 168 GCATCGCGGGGCGCCTTCAGACTAACTGGATCTCTCGGGTTTGTGACAGATGACCTGG 227  
 QY 450 AAATTTATAAATACAAAACGAAAGTGCAGTGGCTCATTTGAAGGACAGCCAAATAGAATA 509  
 DB 228 AAATTTATAAATACAAAACGAAAGTGCAGTGGCTCATTTGAAGGACAGCCAAATAGAATA 287  
 QY 510 GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 569  
 DB 288 GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGGACCATTTATATGTTTATGA 347  
 QY 570 TGGGAGACTCAATTTATGACCGCTAGTTGTGCAATTTAGTGGCCTCATTTGTTCTGAGAG 629  
 DB 348 TGGGAGACTCAATTTATGACCGCTAGTTGTGCAATTTAGTGGCCTCATTTGTTCTGAGAG 407  
 QY 630 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCTTGTGTCGCAATT 689  
 DB 408 AGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCTTGTGTCGCAATT 467  
 QY 690 TTTTAGTCAATGCTCTTATAATTTTGTGAGTTTAAATTTACTTACAGTTTGTGATATGTG 749  
 DB 468 TTTTAGTCAATGCTCTTATAATTTTGTGAGTTTAAATTTACTTACAGTTTGTGATATGTG 527  
 QY 750 TCCAAATTAACCTGCTCAGGCGGAGGAGAGTGTAAAGATCAGTAANTAGCAGCAAACTGTTGA 809  
 DB 528 TCCAAATTAACCTGCTCAGGCGGAGGAGAGTGTAAAGATCAGTAANTAGCAGCAAACTGTTGA 587  
 QY 810 ATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869

Db 588 ATGTGATGTTCTGAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647  
Qy 870 CTGTGGTTCCTCATCGAGGCACTGCAATTCAGATGATGTCAGAGGATGCTCCCTGCTT 929  
Db 648 CTGTGGTTCCTCATCGAGGCACTGCAATTCAGATGATGTCAGAGGATGCTCCCTGCTT 707  
Qy 930 CTGACATGCGAGGGTCTCGATGTTTCAGTTCCTGTACCAAGCTAAACAGTCATTTTGGAC 989  
Db 708 CTGACATGCGAGGGTCTCGATGTTTCAGTTCCTGTACCAAGCTAAACAGTCATTTTGGAC 767  
Qy 990 TCGAGAGGAATATCTAACTTAAGCTCCCGAGAGCATCTCAATAAGCTGTGTCAATGG 1049  
Db 768 TCGAGAGGAATATCTAACTTAAGCTCCCGAGAGCATCTCAATAAGCTGTGTCAATGG 827  
Qy 1050 AAACATTAATGTGGTTCCTGAGAGATATATGTTCAACCACTCAGATTAATAACATGTTCT 1109  
Db 828 AAACATTAATGTGGTTCCTGAGAGATATATGTTCAACCACTCAGATTAATAACATGTTCT 887  
Qy 1110 AGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCGTTCGTGTGAACAATGTGGT 1169  
Db 888 AGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCGTTCGTGTGAACAATGTGGT 947  
Qy 1170 TGTAGATATGTCATTTCTTGGCATTTATACAGGATARAATTTACATGTATCGAGGAAA 1229  
Db 948 TGTAGATATGTCATTTCTTGGCATTTATACAGGATARAATTTACATGTATCGAGGAAA 1007  
Qy 1230 AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTATTTTCAATTCATTAATGAGTC 1289  
Db 1008 AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTATTTTCAATTCATTAATGAGTC 1067  
Qy 1290 ATGGGTGTGTTGACCCCTAAGGCAAGAGAGCAGTATGACGTGGTGGGCACTCTGCACA 1349  
Db 1068 ATGGGTGTGTTGACCCCTAAGGCAAGAGAGCAGTATGACGTGGTGGGCACTCTGCACA 1127  
Qy 1350 CATTTGTACACTGAAGAAATGCGCAGTGGTTCATCTCGTTCATCTTTGGTCACTGCCCTCT 1409  
Db 1128 CATTTGTACACTGAAGAAATGCGCAGTGGTTCATCTCGTTCATCTTTGGTCACTGCCCTCT 1187  
Qy 1410 CTATGATATATAAGCAATGTGAGGAATATGATTTGGATAAGAACACATGAGGATATTT 1469  
Db 1188 CTATGATATATAAGCAATGTGAGGAATATGATTTGGATAAGAACACATGAGGATATTT 1247  
Qy 1470 ACACACCCAGGTGCTTGTGCAAGGGGTTCAGGCCATAGCAGTGTTCACGACCATAG 1529  
Db 1248 ACACACCCAGGTGCTTGTGCAAGGGGTTCAGGCCATAGCAGTGTTCACGACCATAG 1307  
Qy 1530 GACCAGGCGCCCTATACGTTTCATGTGTGCTACAAGGCTTTTCAGTGCCCAATAGTACCGGCT 1589  
Db 1308 GACCAGGCGCCCTATACGTTTCATGTGTGCTACAAGGCTTTTCAGTGCCCAATAGTACCGGCT 1367  
Qy 1590 TGCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGACCAATTCCTTAAGGACAG 1649  
Db 1368 TGCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGACCAATTCCTTAAGGACAG 1427  
Qy 1650 CCGATTTTTCGTTTACTTGCACACAGCTGTGATAGTGAAGTGAACCACTGCTGGTGTGG 1709  
Db 1428 CCGATTTTTCGTTTACTTGCACACAGCTGTGATAGTGAAGTGAACCACTGCTGGTGTGG 1487  
Qy 1710 GGGAAACACACACAAATGACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1769  
Db 1488 GGGAAACACACACAAATGACATCTATGAGCCATGGCGCCAAATGCTTCTCTTCAGATTT 1547  
Qy 1770 CATGGCTATGACATTTGCTGACCGCTGTGATGCTTCCAGACCTGATCTCCACCA 1829  
Db 1548 CATGGCTATGACATTTGCTGACCGCTGTGATGCTTCCAGACCTGAT-TCCACCA 1606  
Qy 1830 TGATGTCAACAGATTTGGCCATT-CAGCAGTCTTTACACAAACAGCACCATTGATGTTCG 1888  
Db 1607 TGATGTCAACAGATTTGGCCATTCCAGCAGTCTTTACACAAACAGCACCATTGATGTTCG 1666  
Qy 1889 GTGGTTTCAATAGTCTCTCTCAGCGACATCTGATTTACCTCGGAACAGTGTGATG 1948  
Db 1667 GTGGTTTCAATAGTCTCTCTCAGCGACATCTGATTTACCTCGGAACAGTGTGATG 1726

Qy 1949 CGCATCGAGTGAAGCCGCTTGTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACA 2008  
Db 1727 CGCATCGAGTGAAGCCGCTTGTGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACA 1786  
Qy 2009 CAGGTCGTCTCAGTGTATCTCGTGGGCGCTGGGCACTGATGAACAAGAAAAAGTTAA 2069  
Db 1787 CAGGTCGTCTCAGTGTATCTCGTGGGCGCTGGGCACTGATGAACAAGAAAAAGTTAA 1846  
Qy 2069 AATCAGATGTTTTCCTCAAAAGAACTCTTGAACCATGACAGATGTACACAGACACAGATT 2128  
Db 1847 AATCAGATGTTTTCCTCAAAAGAACTCTTGAACCATGACAGATGTACACAGACACAGATT 1906  
Qy 2129 GTTACAGCTGTACAGCCCAACCAATGATCTCCACTGGTCAATGACCAATGTGTGCCCA 2189  
Db 1907 GTTACAGCTGTACAGCCCAACCAATGATCTCCACTGGTCAATGACCAATGTGTGCCCA 1966  
Qy 2189 GGAACCAACAGCTGCTCAGAGGCGAGATCTCCATTTTATGATGATGAGAAATTCGCCCAAGG 2248  
Db 1967 GGAACCAACAGCTGCTCAGAGGCGAGATCTCCATTTTATGATGATGAGAAATTCGCCCAAGG 2026  
Qy 2249 ATAAACCCATGTACTGTATTAACAAGAAAGACAGCTGCAGGAGCTGTGCCCTGGACCCAGA 2308  
Db 2027 ATAAACCCATGTACTGTATTAACAAGAAAGACAGCTGCAGGAGCTGTGCCCTGGACCCAGA 2086  
Qy 2309 ACTGCCAGTGGAGGCCCGGGAATCAGGAGTGCAATGGCCCTGCCCGAAAAATATCTGTGGA 2368  
Db 2087 ACTGCCAGTGGAGGCCCGGGAATCAGGAGTGCAATGGCCCTGCCCGAAAAATATCTGTGGA 2146  
Qy 2369 TTGGCTGGCATTTGTTGGAAACTCATGTTTGAATAATTTACTACTGCCAAGGAGAAATATG 2428  
Db 2147 TTGGCTGGCATTTGTTGGAAACTCATGTTTGAATAATTTACTACTGCCAAGGAGAAATATG 2206  
Qy 2429 ACAATGCTAAATGTTCTGTAGGAACCAATGCCCCCTTTTGGCTTCTCTTCAACCCAGA 2489  
Db 2207 ACAATGCTAAATGTTCTGTAGGAACCAATGCCCCCTTTTGGCTTCTCTTCAACCCAGA 2266  
Qy 2489 AGAAGTAGAATTTGTCCTTTAGCAGCTGCGAATATGCAATCATCTCAGACATGTCCA 2548  
Db 2267 AGAAGTAGAATTTGTCCTTTAGCAGCTGCGAATATGCAATCATCTCAGACATGTCCA 2326  
Qy 2549 AGCTCACCTTAACCCCATGGTGGGCTCGGCTTTCGGAAGATCAATGTCTTACTGTCTGGG 2608  
Db 2327 AGCTCACCTTAACCCCATGGTGGGCTCGGCTTTCGGAAGATCAATGTCTTACTGTCTGGG 2386  
Qy 2609 AAGATATGTCCCATTTTACAAATAGTTTACTACAGTGAATGCCCTCTGAGCCCACTGATG 2669  
Db 2387 AAGATATGTCCCATTTTACAAATAGTTTACTACAGTGAATGCCCTCTGAGCCCACTGATG 2446  
Qy 2669 CTGGATTTCTGTGGAATTTTATCAGAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2728  
Db 2447 CTGGATTTCTGTGGAATTTTATCAGAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2506  
Qy 2729 TCAACCCACTCAATGGTGTGTGTGAAGGCCCTGCAAAACACAGTGTCTAAGCAGTGCC 2788  
Db 2507 TCAACCCACTCAATGGTGTGTGTGAAGGCCCTGCAAAACACAGTGTCTAAGCAGTGCC 2566  
Qy 2789 GGAACCATGTGCTTGGAGACAGCATGTGGAGATTTGCAACAGCGGCACTCTGAGTGTCA 2848  
Db 2567 GGAACCATGTGCTTGGAGACAGCATGTGGAGATTTGCAACAGCGGCACTCTGAGTGTCA 2626  
Qy 2849 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTTCCAATGCCCTATGTGGCTCTTCCCTT 2908  
Db 2627 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTTCCAATGCCCTATGTGGCTCTTCCCTT 2686  
Qy 2909 TTGGCAGTGTATGGAAATGATATCAGATGACCACTGCCCGCTTGAANAATGTCAGGCT 2968  
Db 2687 TTGGCAGTGTATGGAAATGATATCAGATGACCACTGCCCGCTTGAANAATGTCAGGCT 2746  
Qy 2969 ACTGTACCTGTAGTCAATTCCTTGGAGCAACAGGCTGTGGCTGGTGTACTGTATCCACGA 3028  
Db 2747 ACTGTACCTGTAGTCAATTCCTTGGAGCAACAGGCTGTGGCTGGTGTACTGTATCCACGA 2806



Db 493 ACGCTTAATACAGCTGTTCCGAAGGTGAGGCTCGAGGAACAGTTCGAGGCAAGCTTC 552  
Qy 411 ----- 410  
Db 553 GGCTACAGATAAGTTCAAGAGTAACTCTGGGGCAACTTGGGCTGTCTCCAAACCAAAA 612  
Qy 411 ----- 410  
Db 613 TGAGCGAAAAGGAGCAAGCTAGAGTCTTTTGGGAAAATTTAGCTGACTAAATTTTTCACC 672  
Qy 411 ----- ACTAATCTGATCTTTGGGGTTTGTGACAGATGGACCTGGAAATTTAATAATACAAA 466  
Db 673 GAGAACTAACTGGCTCTTCTCGGATTTGTAAACAGATGGACCTGGGAAATTTAATAATAAGA 732  
Qy 467 CGAAGTGCAGCTGCTCAATTAAGGAGCAGCAATAAGATAATAGAGCTTCGTTTCAATC 526  
Db 733 CGAAGTGCAGCTGCTCAATTTAAGGAGCAGCAATAAGATAATAGAGCTTCGTTTCAACC 792  
Qy 527 ATTTTGTACAGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGACTCAATTTATG 586  
Db 793 ATTTTGTACAGAAATGTAGCTGGGACCAATTTATATGTTTATGATGGGACTCAATCTAG 852  
Qy 587 CACGGTAGTTGTGCAATTTAGTGGCTCAATGTTCTGAGAGAGATGGCAATGAGACTG 646  
Db 853 CACCTCTGATGCTGCTTTAGTGGCTCAATGTTCTGAAAGAGATGGCAATGAGACGG 912  
Qy 647 TCCCTGAGGTTGTGCGACATCAGGTTATGCTTGTGCAATTTTATGATGCTGCTT 706  
Db 913 CTCTGAGGCTACGTGCTCAATTTAGCTGCTGCAATTTTTCAGTGATGCTGCTT 972  
Qy 707 ATAAATTTGACTGGAATTTAATTTACTTACAGTTTGTGATGTTGTGCTGCAATTTTACTGCTCAG 766  
Db 973 ATAACTGACTGGAATTTAATTTACTTACATTTACATTTTGAATGTTGCTGCAATTTTGTCTCAG 1032  
Qy 767 GCCGAGAGAGTGTAAAGTCAAGTAATAGCAGCAAACTGTTGAATGTAATGTTCTGAAA 826  
Db 1033 GCCGAGAGAGTGTAAAGTCAAGTAATAGCAGCAGCTGTTGAGTGTGAATGTTCTGAAA 1092  
Qy 827 ACTGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAATGTGGTGTTCCTCATC 886  
Db 1093 ACTGAAAGGGGAGTGTGTGACATTCCTCACTGTACAGACAATGTGGTGTTCCTCAOC 1152  
Qy 887 GAGCANTCTCAATTTCAAGTGTGTGAGAGGATGCTCCTGCTCTCAGACTGCGAGGTC 946  
Db 1153 GAGCANTCTGTAATGCAAGGATACCAAGAGGTCCTCCTGCTTCTCATGCGAGGTC 1212  
Qy 947 CTGATGTTCAAGTTCCTGTACAGCTAACAGTCAATTTTGGACTCGAGAGGAATATTCTA 1006  
Db 1213 CTGATGTTCAATTTCTGTGCGAGCTAACAGTCTTTTGGACTCGAGAGGAATATTCTG 1272  
Qy 1007 ACTTAAAGCTCCGAGAGCATCTCAATAGCTGTGTCATGGAACAATTTATGAGGTTG 1066  
Db 1273 ATTTAAAGCTTTCCAGAGCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTTG 1332  
Qy 1067 TTGAGGATATATGTTCAACCACTCAGATTAATACATGTTCTAGCGTATGACCTGCTT 1126  
Db 1333 TTGGCGGATATGTTCAACCACTCAGATTAATACATGTTCTAGCGTATGACCTGACTT 1392  
Qy 1127 CTAGGAGTGGCTTCCATTAACCGTGTGTGAAACAATGTTGTTGTAGATGAGTCAAT 1186  
Db 1393 CTAGGAAATGCTTCCCATTAACCACTGTGTGAAACAATGTTGTTGTAGATGAGTCAAT 1452  
Qy 1187 CTTTGGCATTTACAGAGTAAATTTTACATGTTATGAGGAAATTTGATCCCACTGGGA 1246  
Db 1453 CTTTGGCATTTACAGAGTAAATTTTACATGTTATGAGGAAATTTGATTTCAACAGGA 1512  
Qy 1247 ATGTGACCAATGAGTTGAGAGTTTTCATATTCATATGAGTCAATGAGTGTGTTGACCC 1306  
Db 1513 ACGTGACCAATGAGTGTAGAGTATTTTCAATTCATATGATCAATGAGTATTTGTTAACTC 1572  
Qy 1307 CTAAGGCAAGAGAGATGAGTGTGGGCTCTGCGACATTTGTTACATCTGAAGA 1366

Db 1573 CGAAAGCTTAAGGATCAGTATGCACTGGTGTGACACTCAGCACACATTTGTTACCTGGCAT 1632  
Qy 1367 ATGGCCGAGTGGTCACTGCTGCTCATCTTTGCTCACTGCCCCTCTCTATGGATATATAAGCA 1426  
Db 1633 CTGGCCGTGTGGTCACTGCTGCTCATCTTCGCTCATCTGCCCCTCTATGGATATATAAGCG 1692  
Qy 1427 ATGTGCAGGAATATGATTTGGATTAAGAAACATGAGATATATTACACACCCAGGTTGCC 1486  
Db 1693 TTGTGCAGGAATATGATTTGGAAAGAACACATGAGTATATTACATACTCAGGTTGCTC 1752  
Qy 1487 TTGTGCAGGAGGTTTACGGCCATAGCAGTGTTTACGACCATAGGACAGGCGCCCTATACG 1546  
Db 1753 TTGTGCAGGAGGTTTATGCCCCACAGTAGTGTATTATGATGACAGACCAAGGCTCTGACG 1812  
Qy 1547 TTCTATGTTGGCTTACAAAGCTTTTCAAGTGTGCAATAAGTACCGGCTTGCAGATGATCTTACC 1606  
Db 1813 TTCTATGTTGGCTTACAAAGCTTTTCAAGGCTTTACGGCCCAAAATACCGGCTTGCAGATGACCTCTACA 1872  
Qy 1607 GATATGATGTGGATACCCAGATGTGGACCAATTTCTTAAGCAGACCGGATTTTTCGTTACT 1666  
Db 1873 GATAGATGTGGATACTCAGATGTGGACCAATTTCTTAAGCAGACCGGATTTTTCGTTACT 1932  
Qy 1667 TGCAACACAGCTGTGATAGTGTGAACCATGCTGCTGTTTGGGGGAAACACACAATG 1726  
Db 1933 TGCAACACAGCTGTGATAGTGTGAACCAATGCTGCTGTTTGGAGGGAACACACAATG 1992  
Qy 1727 ACACATCTATAGCCATGCGGCCAAATGCTTCTCTTCAAGATTTTATGCGCTTATGACATG 1786  
Db 1993 ACACATCTATAGCCATGCGGCCAAATGCTTCTCTCTCGGACTTCAATGCTTATGACATG 2052  
Qy 1787 CTTGTGACCGATGCTGCTTCCAGACCTGATCTCCACCATGATCTCAACAGATTTG 1846  
Db 2053 CTTGTGACCGATGCTGCTTCCAGACCTGATCTCCACCATGATCTCAACAGATTTG 2112  
Qy 1847 GCCATTCAGCAGCTTTACAAACAGACCATGATGTGTTGCTGGTGTTCATATGCTCTCC 1906  
Db 2113 GCCATTCAGCAGCTTTTACAAACAGACCATGATGTGTTGCTGGGCTTCAACAGCTCTCC 2172  
Qy 1907 TCCTCAGCGACATCTGTTATTCCTCGGAACAGTGTGATGTCGCTCGGATCGGAGTGAAGCG 1966  
Db 2173 TCCTCAGTACGCTTTGTTTACCTCGGACATGATGTGTTGCTGGGACACACAGTCTCGATGTA 2232  
Qy 1967 CTTGTTTACGACGAGCCTGTTATTCGTTGTGTGGAACACAGGTCGTCTCAGTGTGTA 2026  
Db 2233 CTTGTTGTCAGCAGGACCTGTTATTCGTTGTGTGGAACACAGTCTCGATGTA 2292  
Qy 2027 TCTGTTGGGCTGCGCACTGATGAACAGAAAGTTAAATCAGAACTGTTTTTCCA 2086  
Db 2293 CTTCTGGGAGTTGGCACTGAAGAACAGCAGAAAGTTAAATCAGAGTGTTTTTCTA 2352  
Qy 2087 AAAGAACTCTTGACCATGACATGTGACACGACACACAGATTTGTTACAGCTGTACAGCCA 2146  
Db 2353 AAAGAACTCTTGACCATGACATGTGACACGACACACAGATTTGTTACAGCTGTACAGCCA 2412  
Qy 2147 ACACCAATCACTGCGCACTGGTGCATGACCAATGTTGTGTCGCCAGGAACACAGCTGCTCAG 2206  
Db 2413 ATACCAATCACTGCGCACTGGTGCATGACCAATGTTGTGTCGCCAGGAACACAGCTGCTCAG 2472  
Qy 2207 AAGGCCAGATCTCCATTTTATGATGAAATTTGCCCAAGGATTAACCCCATGTACTACT 2266  
Db 2473 AAGGCCAGATCTCCATTTGCAAGTATGAGTGTGCCCCAAGGATTAACCCCATGTACTACT 2532  
Qy 2267 GTAACAAAGAGACCACTGCTGAGGAGCTGTGCCCTGGAACAGAACTGCCAGTGGAGCCCC 2326  
Db 2533 GCAATTAAGAAACCACTGCTGAGGAGCTGTGCCCTAGACACCACTGCCAGTGGAGCCCC 2592  
Qy 2327 GGNATCAGAGTGCATTTGCCCTGCCGAAAATCTCTGTGGCATTTGCTGGCATTTGGTTG 2386  
Db 2593 GGAATCAAGAGTGCATTTGCCCTGCCGAAAATCTCTGTGGCAATTTGCTGGCATTTGGTTG 2652  
Qy 2387 GAACTCATGTTTGAATAATTTACTTCTGCAAGAGAAATTTATGACATGCTAAATTTGTTCT 2446  
Db 2653 GAACTCATGTTTGAATAATTTACTTCTGCAAGAGAAATTTATGACATGCTAAATTTGTTCT 2712

```
QY 2447 GTAGGAACCAATGCCCTTTTGGCTTCTCTTACAAACCAAGAGGTAGTAATTTGTC 2506
Db 2713 GTAGGAACCAATGCCCTTTTGGCTTCTCTTACAAACCAAGAGGTAGTAATTTGTC 2772
QY 2507 TTAAGCAGCTGCGAATAATGAGTCACTCAGAGCATGTCCAAAGCTCACCTTAACCCCAT 2566
Db 2773 TTAAGCAGCTGCGAATAATGAGTCACTCAGAGCATGTCCAAAGCTCACCTGACTCCAT 2832
QY 2567 GGGTCGGCTTCGGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCCATTTA 2626
Db 2833 GGGTGGCTTCGGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCCATTTA 2892
QY 2627 CAATAGTTTACTACAGTGGATGCCCTGAGCCAGTGTCTGGATTTCTGTGGAATTT 2686
Db 2893 CAATAGTTTCTGAGTGGATGCCCTGAGCCAGTGTCTGGATTTCTGTGGAATTT 2952
QY 2697 TATCAGAAACCCAGTACTCGGGGACTGAAGCTGCAACTGTCATCAACCCACTCAATGGTA 2746
Db 2953 TGTCAAGCTTAGTACTCGGGGATTAAGGCTGCAACTGTCATCAACCCCTCTCAATGGCA 3012
QY 2747 GTGTCTGGAAGGCTGCAACCAACAGTGTCTAGCAGTGCAGACACCATGTGCCCTGA 2806
Db 3013 GCGTCTGGAAGGCTGCAACCAACAGTGTCTAGCAGTGCAGACACCATGTGCCCTGC 3072
QY 2807 GGACAGCATGTGGAGATTGCAACCAAGCGCAGCTCTGAGTGCATGTGGTGCAGCAATGA 2866
Db 3073 GGACAGCATGTGGCAGTGCACATAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGA 3132
QY 2867 AGCAGTGTGGACTCCAAATGCCCTATGTGGCTCTCTCCCTTTTGGCGAGTATGGAAAT 2926
Db 3133 AGCAGTGTGGACTCCAAATGCCCTATGTGGCTCTCTCCCTTTTGGCGAGTATGGAAAT 3192
QY 2927 GGTATACATGACGACCTGCCCCCTGAAATTTGTCAGGCTACTGACCTGTAGTCAAT 2986
Db 3193 GGTATACATGACGAGCTGCCCTGAAATTTGCTCTGGCTACTGTACCTGGAGCAAT 3252
QY 2987 GCTTGGAGCAACAGAGCTGTGGTGTGTAAGTATCCAGCAATATCTGGCAAGGGAAT 3046
Db 3253 GCTTGGAGCAACAGAGCTGTGGTGTGTAAGTATCCAGCAATATCTGGCAAGGGAAT 3312
QY 3047 GCATAGAGGTTTCTATTAAGAGCAGTGAAGATGCTTCGCAAGCCCTACAGGAAT 3106
Db 3313 GTATTGAGGAGCAGTATAAGGACTGTGAAGATGCCGTACAGGCTCTGCGAGGAATG 3372
QY 3107 TCTATCCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATCAACTGGT 3166
Db 3373 TGTATCCACAGCCCTTCTGAATCCAGCATGTGTCTAGAGGACAGCAGATCAACTGGT 3432
QY 3167 CTTTCACTCTGTCAGCTGCAATGCAACGCGCACAGTAAATGCATCAATCAGAGCA 3226
Db 3433 CTTTCACTCTGTCAGCTGTCAGTGCAGTGCAGCAGCAGCAATGCATCAACAGAGTA 3492
QY 3227 TCTGTGAGAGTGTGAGAACTGACCAACAGGCAAGCACTCGAGACTCTGCATATCTGGCT 3286
Db 3493 TCTGTGAGAGTGTGAGAACTGACCAACAGGCAAGCACTCGAGACTCTGCATATCTGGCT 3552
QY 3287 TCTACGCTGATCCCAACCAATGGAGGAAATGTGAGCCATGCAAGTGGCAGCGGT 3346
Db 3553 TCTATGTGACCCGACTTAATGGAGGCAATGTGAGCCATGCAAGTGGCAGCGCAT 3612
QY 3347 CTCCTGTCAACACCAACCAACAGGAGTGTCTTGTGACCAACAGGCGGTCAAGGGGACG 3406
Db 3613 CACTGTGCAACCAACCAACAGGAGTGTCTTGTGACCAACAGGAGTGTCAAGGGGACG 3672
QY 3407 AGTGCCAGCTATGTGAGGTGAAATTCGATACCAAGGAAACCTCTCAGAGGAACATGTT 3466
Db 3673 AGTGCCAGCTATGTGAGGTGAAATTCGATACCAAGGAAACCTCTCAGAGGAACATGCT 3732
QY 3467 ATTATACCTCTTATTGATATCAGTTTACCTTTTATCTATCCAGGAGATGATCGCT 3526
Db 3733 ACTATACCTCTTCAATGATATCAGTTTACCTTTTATCTATCCAGGAGACGACCGCT 3792
```

```
QY 3527 ATTACACAGCTATCAATTTTGGCTACTCTCTGACGAACAAACAGGATTTGGACATGT 3586
Db 3793 ACTACACAGCTATCAATTTTGGCTACTCTCTGATGAACAAACAGGATTTGGACATGT 3852
QY 3587 TCATCAATGCCCTCCAAAGATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTG 3646
Db 3853 TCATCAATGCCCTCCAAACACTTCAACCTCAACATCACCTGGGCTGCCAGTTTCCAGCGG 3912
QY 3647 GAACCCAGCTCGAGAGAGATGCTCTGTTTCAAAAACCAACATTAAGAGATCAAAAG 3706
Db 3913 GAACCCAGCTCGAGAGAGTGCCTGTTTCAAAAACCAACATTAAGAGATCAAAAG 3972
QY 3707 ATAGTTTCTCTAATCAGAAAGTTTGAATTTGCAACACCAAAATATCACTTCTTGT 3766
Db 3973 ATAGTTTCTCTAATGAGAAATTTGAATTTGCAACAGATCCAAACATCACTTCTTGT 4032
QY 3767 ATGTCAATTAATTTCACTGGCCCATCAAAATTCAGATTCCTCTCAGCACAGCAAT 3826
Db 4033 ATGTCAATTAATTTCACTGGCCCATCAAAATTCAGATTCCTCTCAGCACAGCAACT 4092
QY 3827 TTATGACCTGTGACAGTCTCTTCGTGACTTCTTCAGTGTCTCTCTCTCTCTCTCTG 3886
Db 4093 TCATGACCTGTGACAGTCTCTTCGTGACTTCTTCAGTGTCTCTCTCTCTCTCTCTCTG 4152
QY 3887 TGGCTGCTGTGTTTGGAGATCAAAACAAAGTTGTTGGGCTCTCCAGAGCTAGAGAGCAAC 3946
Db 4153 TGGCTGACGTGCTTGGAGATCAAGCAGAGCTGTTGGGCTCTCCAGGCGGAGAGCAAC 4212
QY 3947 TTCTTTCGAGAGTGAACAGATGGCCAGCGCTCCCTTTCCTCTCTCTCTCTCTCTCTG 4006
Db 4213 TTCTTTCGAGAGTGAACAGATGGCCAGCGCTCCCTTTCCTCTCTCTCTCTCTCTCTG 4272
QY 4007 AAACAGATGAGGAGCTCTCTGATCTTATTTGGGGAGTATAAAGACTGTTCCCAAAACCCA 4066
Db 4273 AAACAGATGAGGAGCTCTCTGATCTTATTTGGGGAGTATAAAGACTGTTCCCAAAACCCA 4332
QY 4067 TTGCACTGGAGCGGTGTTTGGCAACAAAGCGCTGTCTCTCTGTGTTGTGAGGTCTC 4126
Db 4333 TTGCACTGGAGCGGTGTTTGGCAACAAAGCGCTGTCTCTCTGTGTTGTGAGGTCTC 4392
QY 4127 CTCGAGGCTGCTGTCATCCCTCTCTGTCGTCAGTCTGCTGTCGTCGTCGTCGTCG 4186
Db 4393 CTCGAGGCTGCTGTCATCCCTCTCTGTCGTCAGTCTGCTGTCGTCGTCGTCGTCG 4452
QY 4187 TGTGCAATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAGTCAAGAGCCGTGAGAA 4246
Db 4453 TGTGCAATTTCTCAGCAGATGCCGATAGTGTACAAGGAGAGTCAAGAGCCGTGAGAA 4512
QY 4247 ACCGGAAGCAGCGCCCTCTGCAAGCTGGAGCTGCTCTGA 4290
Db 4513 ACCGGAAGCAGCGCCCTCTGCAAGCTGGAGCTGCTCTGA 4556
```

## RESULT 12

AAS72660

ID AAS72660 standard; cDNA; 3490 BP.

XX AAS72660;

AC AAS72660;

XX AAS72660;

DT 13-FEB-2002 (first entry)

XX AAS72660;

DE DNA encoding novel human diagnostic protein #8464.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US008631.

XX PF



	XX	31-MAR-2000; 2000US-00540217.
	PR	23-AUG-2000; 2000US-00649167.
	XX	(HYSE-) HYSEQ INC.
	PA	Dmanac RT, Liu C, Tang YT;
	PI	WPI; 2001-639362/73.
	XX	P-PADB; ABG08473.
	DR	New isolated polynucleotide and encoded polypeptides, useful in
	PT	diagnostics, forensics, gene mapping, identification of mutations
	PT	responsible for genetic disorders or other traits and to assess
	PT	biodiversity.
	XX	Claim 1; SEQ ID NO 9464; 103pp; English.
	PS	The invention relates to isolated polynucleotide (I) and polypeptide (II)
	XX	sequences. (I) is useful as hybridisation probes, polymerase chain
	CC	reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
	CC	and in recombinant production of (II). The polynucleotides are also used
	CC	in diagnostics as expressed sequence tags for identifying expressed
	CC	genes. (II) is useful in gene therapy techniques to restore normal
	CC	activity of (II) or to treat disease states involving (II). (II) is
	CC	useful for generating antibodies against it, detecting or quantitating a
	CC	polypeptide in tissue, as molecular weight markers and as a food
	CC	supplement. (II) and its binding partners are useful in medical imaging
	CC	of sites expressing (II). (I) and (II) are useful for treating disorders
	CC	involving aberrant protein expression or biological activity. The
	CC	polypeptide and polynucleotide sequences have applications in
	CC	diagnostics, forensics, gene mapping, identification of mutations
	CC	responsible for genetic disorders or other traits to assess biodiversity
	CC	and to produce other types of data and products dependent on DNA and
	CC	amino acid sequences. AAS64197-AA594564 represent novel human diagnostic
	CC	coding sequences of the invention. Note: the sequence data for this
	CC	patent did not appear in the printed specification, but was obtained in
	CC	electronic format directly from WIPO at
	CC	ftp.wipo.int/pub/published_pct_sequences
	XX	Sequence 3490 BP; 931 A; 772 C; 858 G; 929 T; 0 U; 0 Other;
	SQ	Query Match            65.4%; Score 2807.2; DB 5; Length 3490;
		Best Local Similarity    94.2%; Pred. No. 0;
		Matches 3234; Conservative    0; Mismatches    53; Indels    145; Gaps    25;
Qy	409	AGACTAACTGGATCTTCTCGGTGTTTGACAGATGGCACTGGAAATTATAAATACAAAACG 468
Dd	22	AGACTTACTGGATCTTCTCGGTGTTTGACAGATGGCACTGGAAATTATAAATACAAAACG 81
Qy	469	AAGTCGCATCGTGCTCATTTCAAAGGACGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 528
Dd	82	AAGTCGCATCGTGCTCATTTGAAGGACGCCAAATAGAATAATGAGACTTCGTTTCAATCAT 141
Qy	529	TITGCTACAGATGTAGTTGGACCATTTATATGTTTTATGATGGGACCTCAATTTATGCA 588
Dd	142	TITGCTACAGATGTAGTTGGACCATTTATATGTTTTATGATGGGACCTCAATTTATGCA 201
Qy	589	CCGCTAGTTGCTGCATTTTAGTGGCCTCATTTCTCTGAGAGAGATGGCAATGAGACTGTC 648
Dd	202	CCGCTAGTTGCTGCATTTTAGTGGCCTCATTTCTCTGAGAGAGATGGCAATGAGACTGTC 261
Qy	649	CCTGAGTTGTTGCCACATCAGTTATGCTTGTGCTGCAATTTTTTTTASTGATGCTGCTTAT 708
Dd	262	CCTGAGTTGTTGCCACATCAGTTATGCTTGTGCTGCAATTTTTTTTASTGATGCTGCTTAT 321
Qy	709	AATTTGACTGGATTTAATATTACTTACAGTTTTTCATATGTGCCAAATAACTGCTCAGGC 768
Dd	322	AATTTGACTGGATTTAATATTACTTACAGTTTTTCATATGTGCCAAATAACTGCTCAGGC 381
Qy	769	CGAGGAGAGTGTAAAGATCAGTAATAAGACGAAACTCTTGAATGTGAATGTCTTGAANAAC 828
Dd	382	CGAGGAGAGTGTAAAGATCAGTAATAAGACGCAACTCTTGAATGTGAATGTCTTGAANAAC 441

1906 CTCCTCAGCGACATCCTGGTATTACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCC 1965  
1522 CTCCTCAGCGACATCCTGGTATTACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCC 1581  
1966 GCTTGTGTTAGCAGAGGACCTGGTATTCTGGTGTGTGTGGAAACACAGGGTGTCTCAGTGT 2025  
1582 GCTTGTGTTAGCAGAGGACCTGGTATTCTGGTGTGTGTGGAAACACAGGGTGTCTCAGTGT 1614  
2026 ATCTCGTGGCGCTGGCAACTGATGAACAAGAGAAAAGTTAAATCAGAAATGTTTTTCC 2085  
1615 ----- 1614  
2086 AAAAGAACTCTTGACCATGACAGATGTGACAGACACAGATGTTTACAGCTGPACAGCC 2145  
1615 -----CTCTTTGACCATGACAGATGTGACAGACACAGATGTTTACAGCTGCACAGCC 1667  
2146 AACCCATGACTGCGCACTGGTGAATGCAATGTTGTCTCCCGAGAACCCACAGCTGTCA 2205  
1668 AACCCATGACTGCGCACTGGTGAATGCAATGTTGTCTCCCGAGAACCCACAGCTGTCA 1727  
2206 GAAGCCAGATCTCCATTTTGTAGGTATGAGAAATGCCCCAAGGATAACCCCATGTACTAC 2265  
1728 GAAGGCC-----AGGTATGAGAAATGCCCCAAGGATAACCCCATGTACTAC 1773  
2266 TGTAAACAGAACCCAGCTGACAGAGCTGTGCTCTGACACAGAACTGCCAGTGGAGCCC 2325  
1774 TGTAAACAGAACCCAGCTGACAGAGCTGTGCTCTGACACAGAACTGCCAGTGGAGCCC 1833  
2386 CGGAATCAGGAGTGCAATGCTCCCGAAATATCTGTGCAATGCTGCGCATTTTGGTT 2385  
1834 CGGAATCAGGAGTGCAATGCTCCCGAAATATCTGTGCAATGCTGCGCATTTTGGTT 1893  
2386 GGAACTCATGTTGAAATTTACTGTCGCAAGGAGAAATATGAAATGCTAAATTTGTT 2445  
1894 GGAACTCATGTTGAAATTTACTGTCGCAAGGAGAAATATGAAATGCTAAATTTGTT 1953  
2446 TGTAGGAACCAATGCCCTTTTGGCTCTCTTACAAACCCAGAGAGGTA-GAATTTGT 2504  
1954 TGTAGGAACCAATGCCCTTTTGGCTCTCTTACAAACCCAGAGAGGTAAGGATTTGT 2013  
2505 CTTTAAAG--CAGCTGCGAATATGCA-GTCATCTCAGA-GCATGTCCAA-GCTCACCTTA 2559  
2014 CTTTAAAGCAGCTGCGAATATGCAAGTCACTCAGAGGATGTCAGAGCTCACCTTA 2073  
2560 ACCCATGGG---TCGGCTTCGGAAGATCAA-TGTGTCTACTGTGTGCTGGGAAGATAT 2615  
2074 ACCCATGGGTCGGGCTTCGGGAAGATCAAATTTGTCTACTGTGTGCTGGGAAGATAT 2133  
2616 GTCCCCCATTTACAAATAGTTTACTACAGTGTGATGCCCTCTGAGCCAGTGTGATGCTGAT 2675  
2134 GTCCCCCATTTACAAATAGTTTACTACAGTGTGATGCCCTCTGAGCCAGTGTGATGCTGAT 2193  
2676 CTGTGGAATTTT-----ATCAGAACCCAGTACTCGGGGACTGAAGGTGCAACCTGAT- 2729  
2194 CTGTGGAATTTTATTTCAGGAACCCAGTACTTCGGGGACTGAAGGTGCAACCTGAT 2253  
2730 -CAACCCACTCAATGGTAGTGTCTGTGAAA--GSCCTGCAAAACACAGT-GCTAAGCAGT 2785  
2254 CAACCCACTCAATGGTAGTGTCTGTGAAAAGGCTGTGCAACCAAGTGTGATGCTGAT 2313  
2786 GCCGGACACCATGTGCC-----TTGAGGACAGCATGTGAGATGTGACACCGCAGCTCT 2841  
2314 GCCGGACACCATGTGTGCCCTTTGAGGACAGCATGTGAGATGTGACACCGCAGCTCT 2373  
2842 GAGT-GCATGTGTGTCAGACACATGAA-GCAGTGTGTGACTCCATGCTATGTGCT 2899  
2374 GAGTGGCATGTGTGTCAGACACATGAAAGGAGTGTGTGACTCCATGCTATGTGCT 2433  
2900 CTTTCCCTTTT-GGCCAGTGTATGGAATGTTATACGATGA-GCACCTGCCCCCTGAAAA 2957  
2434 CTTTCCCTTTTGGCCAGTGTATGGAATGTTATACGATGAGGACCTGCCCCCTGAAAA 2493  
2958 TTGTTCCAGGCTACTGTACCTGTAGTCAATGCTGTGAGCAACCA-GGCTGTGGCTGTGTA 3016

2494 TTGTTCCAGCTACTGTACCTGTAGTCAITGCTTGGAGCAACACAGGCTGTGGCTGTGTA 2553  
3017 CTGATCCAGCAATACTGCAAAAGGGAATGATAGAGGTTCTTATTAAGGACCAAGTGA 3076  
2554 CTGATCCAGCAATACTGCAAAAGGGAATGATAGAGGTTCTTATTAAGGACCAAGTGA 2613  
3077 AGATGCTTTGCAAGCCCTCAGGAAATTTCTTATCCACAGCCCTGTCAATTCAGCA 3136  
2614 AGATGCTTTGCAAGCCCTCAGGAAATTTCTTATCCACAGCCCTGTCAATTCAGCA 2673  
3137 TGTGCTAGAGGACAGAGATCAACTG-TCCTTTCTACTGTGCTGAGCTTGCATG 3194  
2674 TGTGCTAGAGGACAGAGATCAACTGCTGCTTTCTTCTACTGTGCTGAGCTTGCATG 2733  
3195 CAAC-GGCCACACAGTAAATGCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCA 3253  
2734 GAAAGGACACAGTAAATGCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCA 2793  
3254 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTAC-GGTGATCCCAACCA-TGGAGG 3311  
2794 CAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGGTGATCCCAACCA-TGGAGG 2853  
3312 GAATGTCAGCCATGCAATGCGGACCGCTCTGTGCAACACACACACGGGCAA 3371  
2854 GAATGTCAGCCATGCAATGCGGACCGCTCTGTGCAACACACACACGGGCAA 2913  
3372 GTGCTTCTGCACCAACCAAGGGCGTCAAGGGGACAGTCCAGCTATGTGAGGTAGAAAA 3431  
2914 GTGCTTCTGCACCAACCAAGGGCGTCAAGGGGACAGTCCAGCTATGTGAGGTAGAAAA 2973  
3432 TCGATACAGGAACCCCTCTCAGAGGAACATGTTATTTATCTCTTCTTATGACTATCA 3491  
2974 TCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTTATCTCTTCTTATGACTATCA 3033  
3492 GTTCACCTTTATGCTATCCAGGAAGATGATCGCTATTATACAGCTATCAATTTTGTGGC 3551  
3034 GTTCACCTTTATGCTATCCAGGAAGATGATCGCTATTATACAGCTATCAATTTTGTGGC 3093  
3552 TACTCTGACGAACAAAAACAGGATTTGGAATGCTATCAATGCTCCCAAGATTTCAA 3611  
3094 TACTCTGACGAACAAAAACAGGATTTGGAATGCTATCAATGCTCCCAAGATTTCAA 3153  
3612 CCTCAACATCAGCTGGCTGCGAGTTCTCAGCTGGAACCCAGGCTGAGAGAGATGCC 3671  
3154 CCTCAACATCAGCTGGCTGCGAGTTCTCAGCTGGAACCCAGGCTGAGAGAGATGCC 3213  
3672 TGTGTTTTCAAAAAACCAACATTAAAGGAGTCAAAAGATAGTTTCTCTAATGAGAAGTTGA 3731  
3214 TGTGTTTTCAAAAAACCAACATTAAAGGAGTCAAAAGATAGTTTCTCTAATGAGAAGTTGA 3273  
3732 TTTTTCGCAACCAACCAATATACATTTCTTTGTTTATGTCAGTAATTTTCACTGGCCAT 3791  
3274 TTTTTCGCAACCAACCAATATACATTTCTTTGTTTATGTCAGTAATTTTCACTGGCCAT 3333  
3792 CAAAATTCAGAT 3803  
3334 CAAAATTCAGG 3345

## RESULT 13

AAZ91922

ID AAZ91922 standard; cDNA; 2625 BP.

XX AC AAZ91922;

XX AC AAZ91922;

DT 08-JUN-2000 (first entry)

XX Human mahogany protein coding sequence #4.

XX Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;  
KW weight regulation; cell therapy; body weight disorder; cachexia;  
KW anorexia; hyperpigmentation; increased metabolic rate disorder;

KW hyperphagia; Antibesity; antianorexic; anticachexic; ss.  
 XX Homo sapiens.  
 XX WO200005373-A2.  
 XX 03-FEB-2000.  
 XX 21-JUL-1999; 99WO-US016484.  
 XX 21-JUL-1998; 98US-0093630P.  
 PR 20-OCT-1998; 98US-0104978P.  
 PR 05-FEB-1999; 99US-00245041.  
 XX (MILL-) MILLINIUM PHARM INC.  
 XX Moore K, Nagle DL;  
 XX WPI: 2000-195103/17.  
 DR P-PSDB; AAY81809.  
 XX New human and murine mahogany genes, useful, e.g. for diagnosis and  
 PT treatment of body weight disorders.  
 XX Claim 1; Fig 20a; 188pp; English.  
 XX This sequence represents a human mahogany gene of the invention. The  
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iii) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in (cell-  
 CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (II) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (II) (for diagnosis, prognosis or monitoring of treatment); to  
 CC evaluate (II)-expressing cells intended for cell therapy, and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity. (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other  
 CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate  
 XX Sequence. 2625 BP; 660 A; 605 C; 694 G; 666 T; 0 U; 0 Other;

Query Match 52.4%; Score 2248.2; DB 3; Length 2625;  
 Best Local Similarity 96.8%; Pred. No. 0;  
 Matches 2308; Conservative 0; Mismatches 68; Indels 9; Gaps 1;

QY 3 GGTGCGCGCAGCGCGCACTGAGGCAAGCTCAGGAGGAGGACGGCGCGCAGCGCAGC 62  
 Db 9 GGTGCGCGCAGCGCGCGCACTGAGGCAAGCTCAGGAGGAGGAGGAGGCGCGCGCAGC 68  
 QY 63 GCTCGCGCGCAGGAGCG 122  
 Db 69 GCTCGCGCGCAGGAGCG 128  
 QY 123 GGGGCTGGGCG 182  
 Db 129 GGGGCG 188  
 QY 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242  
 Db 189 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 239  
 QY 243 CGAGCG 302  
 Db 240 CGAGCG 299  
 QY 303 TGACCG 362

Db 300 TGACCGCGCGGTGTCAACGGCGGTGCTGCTGCAACCTGGCACCGGCGAGTGGCTGCGCC 359  
 QY 363 CGCGCGCTGGGTGGCGGAGCAATGCGACACTGCGGGGGCGCGCTTCAGACTAACTGGATC 422  
 Db 360 CGCGCGCTGGGTGGCGGAGCAATGCGACACTGCGGGGGCGCGCTTCAGACTAACTGGATC 419  
 QY 423 TTCTGGGTTTGTGACAGATGGACCTGGAAATATATAATACAAAAGAGTGCACCTGGCT 482  
 Db 420 TTCTGGGTTTGTGACAGATGGACCTGGAAATATATAATACAAAAGAGTGCACCTGGCT 479  
 QY 483 CATTTGAAGGACAGCCCAATAGAAATAGAGACTTCGTTTCAATCATTTTGTACAGATG 542  
 Db 480 CATTTGAAGGACAGCCCAATAGAAATAGAGACTTCGTTTCAATCATTTTGTACAGATG 539  
 QY 543 TAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC 602  
 Db 540 TAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC 599  
 QY 603 ATTTAGTGGCTCATTTGCTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGC 662  
 Db 600 ATTTAGTGGCTCATTTGCTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGC 659  
 QY 663 CACATCAGGTTATGCTTGTGCTCAATTTTATGATGCTGCTTATAATTTGACTGGATT 722  
 Db 660 CACATCAGGTTATGCTTGTGCTCAATTTTATGATGCTGCTTATAATTTGACTGGATT 719  
 QY 723 TAATATTACTTACAGTTTGTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCT 782  
 Db 720 TAATATTACTTACAGTTTGTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCT 779  
 QY 783 GATCAGTAAATAGCAGCAAACTGTTGAAATGTAATGTTTCTGAAAACCTGAAAAGTGAAGC 842  
 Db 780 GATCAGTAAATAGCAGCAAACTGTTGAAATGTAATGTTTCTGAAAACCTGAAAAGTGAAGC 839  
 QY 843 ATGTGACATTCCTCTACTGTACAGCAACTGTGTTTCTCTCATCGAGCACTGCAATTC 902  
 Db 840 ATGTGACATTCCTCTACTGTACAGCAACTGTGTTTCTCTCATCGAGCACTGCAATTC 899  
 QY 903 AAGTGAATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAGGGTCTCGGATGTTCACTTCC 962  
 Db 900 AAGTGAATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAGGGTCTCGGATGTTCACTTCC 959  
 QY 963 TGTACAGCTAACAGTCAATTTTGGACTGCGAGAGAAATTTCTAACTTAAAGCTCCCGAC 1022  
 Db 960 TGTACAGCTAACAGTCAATTTTGGACTGCGAGAGAAATTTCTAACTTAAAGCTCCCGAC 1019  
 QY 1023 AGCATCTCAATGAAGCTGTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCT 1082  
 Db 1020 AGCATCTCAATGAAGCTGTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCTGTCCT 1079  
 QY 1083 CAACCACTCAGATTATAACATGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCC 1142  
 Db 1080 CAACCACTCAGATTATAACATGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCC 1139  
 QY 1143 ACTAAACCGTTCTGTGAACAAATGTTGTTAGATATGTTTCTTCTTGGCAATTATACAA 1202  
 Db 1140 ACTAAACCGTTCTGTGAACAAATGTTGTTAGATATGTTTCTTCTTGGCAATTATACAA 1199  
 QY 1203 CGATAAAATTTACATGTAAGGAGGAAATTCATCCAACTGGGAAATGTGACCAATGAGTT 1262  
 Db 1200 CGATAAAATTTACATGTAAGGAGGAAATTCATCCAACTGGGAAATGTGACCAATGAGTT 1259  
 QY 1263 GAGAGTTTTCATTCATTAATGAGTCATGGGTGTTGTTGACCCCTTAAGGCAAGGAGCA 1322  
 Db 1260 GAGAGTTTTCATTCATTAATGAGTCATGGGTGTTGTTGACCCCTTAAGGCAAGGAGCA 1319  
 QY 1323 GTATGAGTGGTTGGGCACTCTGCACACATTTGTTACCTGGAAGAATGGCCGAGTGGTCA 1382  
 Db 1320 GTATGAGTGGTTGGGCACTCTGCACACATTTGTTACCTGGAAGAATGGCCGAGTGGTCA 1379  
 QY 1383 GCTGCTCATCTTGTGCTCACTGCGCTCTCTATGATATATAAGCAATGTGCGAGAAATAGA 1442  
 Db 1380 GCTGCTCATCTTGTGCTCACTGCGCTCTCTATGATATATAAGCAATGTGCGAGAAATAGA 1439







Qy 1187 CTTTGGCAATTATACAGGATAAATTTACATGTATGGAGAAAATTTGATCCAACTGGGA 1246  
Db 1453 CTTTGGCAATTATACAGGATAAATTTACATGTATGGAGAAAATTTGATTCACAGGGA 1512  
Qy 1247 ATGTGACCAATGAGTTCAGAGTATTTTTCACATTCATATAGTCAATGGGTGTTGACCC 1306  
Db 1513 ACGTGACCAATGAGTTCAGAGTATTTTTCATATTCATATAGTCAATGGGTGTTAACTC 1572  
Qy 1307 CTAAGGCAAGAGGAGCATGTCAGTGTGGGCACTCTGCACACATTTGTACACTGAAGA 1366  
Db 1573 CGAAAGCTAAGGATCAGTATGAGTGTGGGCACTCAGCACACATTTGTACACTGGCAT 1632  
Qy 1367 ATGGCCGAGTGTTCATCTGTCATCTTTGTCCTGCTCTCTATGATATATTAAGCA 1426  
Db 1633 CTGGCCGAGTGTTCATCTGTCATCTTTGTCCTGCTCTCTATGATATATTAAGCG 1692  
Qy 1427 ATGTGACGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCC 1486  
Db 1693 TTGTGACGGAATATGATTTGGATAAGAACACATGGAGTATATTACATACCTCAGGGTGCTC 1752  
Qy 1487 TTGTGCAAGGGGTTACGGCCATAGCAGTGTTCAGCACCATAGGACCGAGGGCCCTATACG 1546  
Db 1753 TTGTGCAAGGGGTTATGGCCACAGTAGTGTTCATGATGACAGCACCGAGGCTCTGTACG 1812  
Qy 1547 TTCAATGTTGGCTCAAGGCTTTTCAAGTCCCAATAGTACCGGCTTGCAGATGATCTTACC 1606  
Db 1813 TTCAATGTTGGCTCAAGGCTTTTCAAGTCCCAATAGTACCGGCTTGCAGATGATCTTACA 1872  
Qy 1607 GATATGATGTGGATACCCAGATGTGGACCATCTTAAAGACAGCCGATTTTTCGGTTACT 1666  
Db 1873 GATACGATGTGGATACCTCAGATGTGGACCATCTTAAAGACAGCCGATTTTTCGGTTACT 1932  
Qy 1667 TGCAACAGCTGTGATAGTGTGAGTGGAAACCATGCTGCTGTTTGGGGGAAACACACAAATG 1726  
Db 1933 TGCAACAGCTGTGATAGTGTGAGTGGAAACCATGCTGCTGTTTGGGGGAAACACACAAATG 1992  
Qy 1727 ACACATCTATGAGCATGGGCCAAATGCTTCTCTTTCAGATTTTCAATGGCCCTATGCAATTG 1786  
Db 1993 ACACATCTATGAGCATGGGCCAAATGCTTCTCTTTCAGATTTTCAATGGCCCTATGCAATTG 2052  
Qy 1787 CCTGTGACCGTGTGATGCTTCCAGACATGATCTCCACCATGATGTCACAGATTTTG 1846  
Db 2053 CTTGTGACCGATGATGCTTCCAGACATGATCTCCACCATGATGTCACAGATTTTG 2112  
Qy 1847 GCCATTGACAGTCTTACACACAGCACCATGATGTTGCGGTGTTTCAATAGTCTCC 1906  
Db 2113 GCCATTGACAGTCTTACACACAGCACCATGATGTTGCGGTGTTTCAATAGTCTCC 2172  
Qy 1907 TCCTCAGGACATCTGTGATTTACCTCGGACAGTGTGATGCGCATCGGAGTGAAGCG 1966  
Db 2173 TCCTCAGTACGCTCTTGTCTTTTACCTCGGAGCAGTGTGATGCGCATCGGAGTGAAGCTG 2232  
Qy 1967 CTTGTTTACGACGAGGACCTGATTTGCTGTGTGTGGAACACAGGCTGCTCAGTGTA 2026  
Db 2233 CTTGTTTACGACGAGGACCTGATTTGCTGTGTGTGGAACACAGGCTGCTCAGTGTA 2292  
Qy 2027 TCTCGTGGGCTGGCACTGATGAACAGAGAAAGTTAAATCAGATGTTTTTCCA 2086  
Db 2293 CTTCTGGGAGTGGCACTGATGAACAGAGAAAGTTAAATCAGATGTTTTTCCA 2352  
Qy 2087 AAGGACCTTTGACCATGACAGATGTGACAGCACAGATTTGTACAGCTGTACAGCCA 2146  
Db 2353 AAGGACCTTTGACCATGACAGATGTGACAGCACAGATTTGTACAGCTGTACAGCCA 2412  
Qy 2147 ACACCAA 2153  
Db 2413 ATACCAA 2419

Search completed: March 4, 2004, 18:09:30  
Job time : 1030.9 secs



**THE ROYAL CANADIAN MOUNTED POLICE**





Db	2280	CTGCAGAGAGCTGTGCCCTTGAGCCAGAACTGCCAGTGGGAGGCCCGGAATCAGAGATGCAT	2333
Qy	2343	TGCCCTGCCCGAAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAAACTCATGTTTGAA	2402
Db	2340	TGCCCTGCCCGAAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAAACTCATGTTTGAA	2399
Qy	2403	AATTACTACTGCCAAGGAGAAATTATGCAANTGCTAAATTTGTTCTGTGTAGGAAACACAATGC	2462
Db	2400	AATTACTACTGCCAAGGAGAAATTATGCAANTGCTAAATTTGTTCTGTGTAGGAAACACAATGC	2459
Qy	2463	CCTTTTGGCTTCTCTTACAAACCCAGAAAGGTAGAAATTTGTCTTTAAAGCAGCTGCGAAT	2522
Db	2460	CCTTTTGGCTTCTCTTACAAACCCAGAAAGGTAGAAATTTGTCTTTAAAGCAGCTGCGAAT	2519
Qy	2523	AATGCAGTCACTCAGAGCATGTGCCAAGCTCACTTAAACCCATGGGTGGCCCTTCGGAA	2582
Db	2520	AATGCAGTCACTCAGAGCATGTGCCAAGCTCACTTAAACCCATGGGTGGCCCTTCGGAA	2579
Qy	2583	GATCAATGTCTCTACTGTGTCTGGGAAGATATGTCCCACTTTACAAATAGTTTACTACA	2642
Db	2580	GATCAATGTCTCTACTGTGTCTGGGAAGATATGTCCCACTTTACAAATAGTTTACTACA	2639
Qy	2643	GTGGATGCCGTCTGAGCCCATGATCTGGAATTTCTGTGGAAATTTTATCAGAACCCAGTAC	2702
Db	2640	GTGGATGCCGTCTGAGCCCATGATCTGGAATTTCTGTGGAAATTTTATCAGAACCCAGTAC	2699
Qy	2703	TGCGGACTCAAGGCTGCAACTGCATCAACCCACCACTCAATGSGTAGTGTCTGTGCAAGGCC	2762
Db	2700	TGCGGACTCAAGGCTGCAACTGCATCAACCCACCACTCAATGSGTAGTGTCTGTGCAAGGCC	2759
Qy	2763	TGCAAAACCAAGTCTAAGCAGTCCCGGACACCATGTGCCCTTTGAGGACAGCATGTGGAGA	2822
Db	2760	TGCAAAACCAAGTCTAAGCAGTCCCGGACACCATGTGCCCTTTGAGGACAGCATGTGGAGA	2819
Qy	2823	TTGCACAGCGGCAGCTCTGAGTGCACTGGTGCAAGAAATGATGAGCAAGTGTGTGCACTC	2882
Db	2820	TTGCACAGCGGCAGCTCTGAGTGCACTGGTGCAAGAAATGATGAGCAAGTGTGTGCACTC	2879
Qy	2883	CAATGCCCTATGTGGCCCTCCCTTCCCTTTTGCCAGTGTATGGAATGSGTATACGATGAGCAC	2942
Db	2880	CAATGCCCTATGTGGCCCTCCCTTCCCTTTTGCCAGTGTATGGAATGSGTATACGATGAGCAC	2939
Qy	2943	CTGCCCCCTGAAAATTTGTCAGGCTACTGTATCTGTAGTCACTGTCTTGAGCAACACAGG	3002
Db	2940	CTGCCCCCTGAAAATTTGTCAGGCTACTGTATCTGTAGTCACTGTCTTGAGCAACACAGG	2999
Qy	3003	CTGTGGCTGTGTACTCATCCCAAGCAATCTGGCAAAAGGAAATGCAATACAGGGTTCCCTA	3062
Db	3000	CTGTGGCTGTGTACTCATCCAGCAATCTGGCAAAAGGAAATGCAATACAGGGTTCCCTA	3059
Qy	3063	TAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCAAGGCCCT	3122
Db	3060	TAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCAAGGCCCT	3119
Qy	3123	GCTCAATTCAGCATGTGTTAGAGGACAGCAGATCAACTGTGCTTTCAATTCACATGTCC	3182
Db	3120	GCTCAATTCAGCATGTGTTAGAGGACAGCAGATCAACTGTGCTTTCAATTCACATGTCC	3179
Qy	3183	AGCTTGCCAAATGCAACGCCACACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGA	3242
Db	3180	AGCTTGCCAAATGCAACGCCACACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGA	3239
Qy	3243	GAACTGTGACCAAGGCAAGCATCTCGGAGACTGTGATATCTGGCTTCTACCGTGTATCCAC	3302
Db	3240	GAACTGTGACCAAGGCAAGCATCTCGGAGACTGTGATATCTGGCTTCTACCGTGTATCCAC	3299
Qy	3303	CAATGTGAGGGAATATGTGAGCCATGCAAGTGAATGGGCACGCGTCTCTGTGCAACACCAA	3362
Db	3300	CAATGTGAGGGAATATGTGAGCCATGCAAGTGAATGGGCACGCGTCTCTGTGCAACACCAA	3359
Qy	3363	CACGGGCAAGTGTCTTGTGACCAACAAAGGGCGTCAAGGGGAGCAGGTGCCAGTATGTGA	3422
Db	3360	CACGGGCAAGTGTCTTGTGACCAACAAAGGGCGTCAAGGGGAGCAGGTGCCAGTATGTGA	3419

QY	3423	GGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGACATGTTATTATTA	CTTCTCTTAT	3482
Db	3420	GGTAGAAATTCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTATTA	CTTCTCTTAT	3479
QY	3483	TGACTATCAGTTCCACCTTTAGTCTATCCCAAGGAAGATGATCGCTATTATCACAGCTATCAA	3542	
Db	3480	TGACTATCAGTTCCACCTTTAGTCTATCCCAAGGAAGATGATCGCTATTATCACAGCTATCAA	3539	
QY	3543	TTTTTGGGTACTCTCGACGAAACAAACACGGGATTTGGACATGTCATCAATGCTCCAA	3602	
Db	3540	TTTTTGGGTACTCTCGACGAAACAAACACGGGATTTGGACATGTCATCAATGCTCCAA	3599	
QY	3603	GAATTTCAACCTCAACATCACCTCGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGA	3662	
Db	3600	GAATTTCAACCTCAACATCACCTCGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGA	3659	
QY	3663	AGAGATGCCCTGTTGTTTCAAAAAACCAACATTAAGGAGTACAAGATAGTTTTCTCTAATGA	3722	3722
Db	3660	AGAGATGCCCTGTTGTTTCAAAAAACCAACATTAAGGAGTACAAGATAGTTTTCTCTAATGA	3719	
QY	3723	GAAGTTTGATTTTCGCAACCAACCAATATCACCTTTCTTTGTTTATGTCAGTAATTTTCC	3782	
Db	3720	GAAGTTTGATTTTCGCAACCAACCAATATCACCTTTCTTTGTTTATGTCAGTAATTTTCC	3779	
QY	3783	CTGGCCCATCAAAATTCAGATTCAGATTCGCTCTCTCAGCACAGCAATTTTATGACCTGGTACA	3842	
Db	3780	CTGGCCCATCAAAATTCAGATTCAGATTCGCTCTCTCAGCACAGCAATTTTATGACCTGGTACA	3839	
QY	3843	GTTCCTTCGTCACATTTCTTTTCAGTTGTTTTCCTCTCTTTGCTCTCGTGGCTGCTGTGGTTTG	3902	
Db	3840	GTTCCTTCGTCACATTTCTTTTCAGTTGTTTTCCTCTCTTTGCTCTCGTGGCTGCTGTGGTTTG	3899	
QY	3903	GAAGATCAAAACAAAGTTGTTGGGCTCCAGACGTAGAGAGCACTTCTTCGAGAGATGCA	3962	
Db	3900	GAAGATCAAAACAAAGTTGTTGGGCTCCAGACGTAGAGAGCACTTCTTCGAGAGATGCA	3959	
QY	3963	ACAGATGGCCAGCGCTCCCTTTGCTCTGTAATGTCGCCCTTGAAACAGATGAGAGCC	4022	4022
Db	3960	ACAGATGGCCAGCGCTCCCTTTGCTCTGTAATGTCGCCCTTGAAACAGATGAGAGCC	4019	
QY	4023	TCCTGATCTTATTTGGGGGGAGTATAAGACTGTTCACAAACCAATTCGACTGGAGCCGTG	4082	
Db	4020	TCCTGATCTTATTTGGGGGGAGTATAAGACTGTTCACAAACCAATTCGACTGGAGCCGTG	4079	
QY	4083	TTTTTGGCAACAAAGCCCTGCTCTCTGTTGTTGTGAGGCTCCCTCGAGGCTTCGGTGG	4142	
Db	4080	TTTTTGGCAACAAAGCCCTGCTCTCTGTTGTTGTGAGGCTCCCTCGAGGCTTCGGTGG	4139	
QY	4143	CATCCCTCTCTGGGAGTCAGTCTTTGCTGTGGCCAGCGCCCTGTTGNCATTTCTCA	4202	
Db	4140	CATCCCTCTCTGGGAGTCAGTCTTTGCTGTGGCCAGCGCCCTGTTGNCATTTCTCA	4199	
QY	4203	GCAGATGCCGATAGTGTACAAAGGAAGAAGTCAGGAGCCGTGAGAAACCGGAGACGAGCC	4262	
Db	4200	GCAGATGCCGATAGTGTACAAAGGAAGAAGTCAGGAGCCGTGAGAAACCGGAGACGAGCC	4259	
QY	4263	CCCTTGCAACGCTGGGACCTGCATCTGA	4290	
Db	4260	CCCTTGCAACGCTGGGACCTGCATCTGA	4287	

RESULT 2  
US-09-893-238-16  
; Sequence 16, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238

```

, CURRENT FILING DATE: 2001-06-27
, PRIOR APPLICATION NUMBER: 09/245,041
, PRIOR FILING DATE: 1999-02-05
, PRIOR APPLICATION NUMBER: 60/093,630
, PRIOR FILING DATE: 1998-07-21
, PRIOR APPLICATION NUMBER: 60/104,978
, PRIOR FILING DATE: 1998-10-20
, NUMBER OF SEQ ID NOS: 129
, SOFTWARE: FastSeq for Windows Version 3.0
, SEQ ID NO 16
, LENGTH: 4072
, TYPE: DNA
, ORGANISM: Homo sapiens
US-09-893-238-16

```

Query Match	86.2%	Score 3696.2	DB 9	Length 4072
Best Local Similarity	98.5%	Pred. No. 0		
Matches 3744	Conservative 0	Mismatches 48	Indels 9	Gaps 1
Qy	3	GGTGGCCGACGGCGGCAACTGAGGCAAGGCTGAGAGGAGGACGGCGGACGGCAGC	62	
Db	9	GGTGGCCGACGGCGGCAACTGAGGCAAGGCTGAGAGGAGGACGGCGGACGGCAGC	68	
Qy	63	GCTCGGGGACGAGAGCGCGGGCCGCACTGGGA CTGGGA CTGACCGAGGGCTGGAGGCC	122	
Db	69	GCTCGGGGACGAGAGCGCGGGCCGCACTGGGA CTGGGA CTGACCGAGGGCTGGAGGCC	128	
Qy	123	GGGCTCGGGGGCGGGGTGCGCCTCCCGCGCTGCTCTCCACCGCTCGGGCCACCGCT	182	
Db	129	GGGACCGCGCCCGGTGTGTCTCCCGGGGTCTGTGGGGGCGCTGCCCGCGCGCC	188	
Qy	183	GCTGCTCTCTGTGTGTGTGCTCCCGCGCCGCTGTGCTGCTGTGCTGCCCTGTGAGGC	242	
Db	189	GCTGCTCCGCTGTCTT-----TTGCTGCTGCTGCTGCGGCTGCCCGGAGGC	239	
Qy	243	CGAGGCGCGCGCGCGCGCGGGCGGTGTCGGGCTCAGCCGACCGCAGGCCACGAGGATG	302	
Db	240	CGAGGCGCGCTGCGGTGGCGCGGGGTGTCCGGCTCGGCGCGAGCCACGAGGATG	299	
Qy	303	TGACCGGCCCTGTGTCAACGGCGGTCTGCAACCCCTGGCACCGCGCAGTGCCTGTGCC	362	
Db	300	TGACCGGCCCTGTGTCAACGGCGGTCTGCAACCCCTGGCACCGCGCAGTGCCTGTGCC	359	
Qy	363	CGCGGCTGAGTGGGCGAGCAATGCCAGCACTGGGGGGCGGCTTCAGACTTAATGGATC	422	
Db	360	CGCGGCTGAGTGGGCGAGCAATGCCAGCACTGGGGGGCGGCTTCAGACTTAATGGATC	419	
Qy	423	TTCTGGGTTTGTGACAGATGGACCTGGAAATTATAATACAAACGAAAGTCACGTGGCT	482	
Db	420	TTCTGGGTTTGTGACAGATGGACCTGGAAATTATAATACAAACGAAAGTCACGTGGCT	479	
Qy	483	CATTGAGGACGCGCAATAGATTAATGACACTTCGTTTCAATCACTTTGCTACAGATG	542	
Db	480	CATTGAGGACGCGCAATAGATTAATGACACTTCGTTTCAATCACTTTGCTACAGATG	539	
Qy	543	TAGTTGGGACCATTTATGTTTATGATGGGACTCAATTTATGCAACCGCTAGTTGCTGC	602	
Db	540	TAGTTGGGACCATTTATGTTTATGATGGGACTCAATTTATGCAACCGCTAGTTGCTGC	599	
Qy	603	ATTAGTGGGCTCATTTGCTCGAGAGAGATGGCAATGACACTGTCCTCAGGTTGTC	662	
Db	600	ATTAGTGGGCTCATTTGCTCGAGAGAGATGGCAATGACACTGTCCTCAGGTTGTC	659	
Qy	663	CACATCAGGTTATGCCCTTGCTGATTTTTTTAGTGTGCTGTATAATTTGACTGGATT	722	
Db	660	CACATCAGGTTATGCCCTTGCTGATTTTTTTAGTGTGCTGTATAATTTGACTGGATT	719	
Qy	723	TAAATATTACATTACAGTTTTCATATGTCACAAATTAATGCTCAGGCGGAGGAGTGTA	782	
Db	720	TAAATATTACATTACAGTTTTCATATGTCACAAATTAATGCTCAGGCGGAGGAGTGTA	779	
Qy	783	GATCAGTAAATAGCAGCGAAACTGTGTAATGTGAATGTTCTGAAACATGGAAGGTGAAGC	842	

QY	1923	GGTATTCACTCGGAACAGTGTGATCGGCATCGGAGTGAAGCGCTTGTTTATGACGACAG	1982
DB	1920	GGTATTCACTCGGAACAGTGTGATCGGCATCGGAGTGAAGCGCTTGTTTATGACGACAG	1979
QY	1983	ACCTGGTATTTCGGTGTGTGTGGAAACAAGGGTTCGTCTCAGTGTATCTCGTGGCGCGTGC	2042
DB	1980	ACCTGGTATTTCGGTGTGTGTGGAAACAAGGGTTCGTCTCAGTGTATCTCGTGGCGCGTGC	2039
QY	2043	AACTGATGAACAAGAAGAAAAGTTAAATACAGAAATGTTTTCCAAAAGAACTCTTGACCA	2102
DB	2040	AACTGATGAACAAGAAGAAAAGTTAAATACAGAAATGTTTTCCAAAAGAACTCTTGACCA	2099
QY	2103	TGACAGANTGACACAGCAACAGATTGTTACAGCTGTACAGCCAAACACCAATGACTGGCA	2162
DB	2100	TGACAGANTGACACAGCAACAGATTGTTACAGCTGTACAGCCAAACACCAATGACTGGCA	2159
QY	2163	CTGGTGCAAATGACCATTTGTCCTCCCAAGGAACCAACAGCTGCTCAGAAAGGCCAGATCTCCAT	2222
DB	2160	CTGGTGCAAATGACCATTTGTCCTCCCAAGGAACCAACAGCTGCTCAGAAAGGCCAGATCTCCAT	2219
QY	2223	TTTTAGGTATGAGAAATTGCCCCCAAGGATTAACCCCATGTACTACTGTAAACAGNAGCCAG	2282
DB	2220	TTTTAGGTATGAGAAATTGCCCCCAAGGATTAACCCCATGTACTACTGTAAACAGNAGCCAG	2279
QY	2283	CTGCAGAGCTGTGCCCTGGACACAGAACTGCCAGTGGAGCCCCGGAAATCAGAGTGCAT	2342
DB	2280	CTGCAGAGCTGTGCCCTGGACACAGAACTGCCAGTGGAGCCCCGGAAATCAGAGTGCAT	2339
QY	2343	TGCCCTGCCGAAAAATATCTGTGGCATTTGCTGGCATTTGGTTTGGAAATCATCTGTTTGA	2402
DB	2340	TGCCCTGCCGAAAAATATCTGTGGCATTTGCTGGCATTTGGTTTGGAAATCATCTGTTTGA	2399
QY	2403	AAATTACTACTGCCAAGGAGAAATTATGACAAATGCTAAATTTCTCTGTAGGAACACCAATGC	2462
DB	2400	AAATTACTACTGCCAAGGAGAAATTATGACAAATGCTAAATTTCTCTGTAGGAACACCAATGC	2459
QY	2463	CTTTTGGCTTCTCTTACAAACCCAGAGAAGTAGAAATTTGTCCTTAAGCAGCTGGCAAT	2522
DB	2460	CTTTTGGCTTCTCTTACAAACCCAGAGAAGTAGAAATTTGTCCTTAAGCAGCTGGCAAT	2519
QY	2523	AATGCACTCATCTCAGAGCATGTCACAGCTCACCTTAAACCCCATGGCTGGCCCTTCGGAA	2582
DB	2520	AATGCACTCATCTCAGAGCATGTCACAGCTCACCTTAAACCCCATGGCTGGCCCTTCGGAA	2579
QY	2583	GATCAATGTGTCTTACTGTTGGTGTGGGAAGATATGTCCCCATTTTACAAATAGTTTACTACA	2642
DB	2580	GATCAATGTGTCTTACTGTTGGTGTGGGAAGATATGTCCCCATTTTACAAATAGTTTACTACA	2639
QY	2643	GTGATGCGGTCTCAGCCGACAGTGTGATGCTGGATTCCTGTGGAAATTTTATACAAACCCAGTAC	2702
DB	2640	GTGATGCGGTCTCAGCCGACAGTGTGATGCTGGATTCCTGTGGAAATTTTATACAAACCCAGTAC	2699
QY	2703	TCGGGGACTGAAGGCTGCAACCTCCATCAACCCCACTCAATGGTAGTGTCTGTGAAAGGCC	2762
DB	2700	TCGGGGACTGAAGGCTGCAACCTCCATCAACCCCACTCAATGGTAGTGTCTGTGAAAGGCC	2759
QY	2763	TGCAAAACACAGTGTCTAAGCAGTCCCGACACCATGTGCCCTTGGAGGACAGCATGTGGAGA	2822
DB	2760	TGCAAAACACAGTGTCTAAGCAGTCCCGACACCATGTGCCCTTGGAGGACAGCATGTGGAGA	2819
QY	2823	TTGACACAGCGGACGCTAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTC	2882
DB	2820	TTGACACAGCGGACGCTAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTC	2879
QY	2883	CAATGSCCTATGTGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGACAC	2942
DB	2880	CAATGSCCTATGTGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGACAC	2939
QY	2943	CTGCCCCCTGAAAAATTTGTTCCAGGCTACTGTACTCTGTAGTCAITTCGTTGGAGCAACGAG	3002
DB	2940	CTGCCCCCTGAAAAATTTGTTCCAGGCTACTGTACTCTGTAGTCAITTCGTTGGAGCAACGAG	2999

QY	3003	CTGTGGCTGGTGTA	CTGATCCAGCAAT	ACTTGGCAAAAGGAAAT	TGCATAGAGGGTTCCTA	3067
DB	3000	CTGTGGCTGGTGTA	CTGATCCAGCAAT	ACTGCGAAAGGAAAT	TGCATAGAGGGTTCCTA	3059
QY	3063	TAAAGACCACTG	GAAGATGCTTCGCAAGCCCT	TACAGGAAATTTCTATCCACAGCCCT	3122	
DB	3060	TAAAGACCACTG	GAAGATGCTTCGCAAGCCCT	TACAGGAAATTTCTATCCACAGCCCT	3119	
QY	3123	GCTCAATTC	CAGCATGTGTAGAGGACAGCAGATACA	CTGGTCTTTCAATTCATCTGTCC	3182	
DB	3120	GCTCAATTC	CAGCATGTGTAGAGGACAGCAGATACA	CTGGTCTTTCAATTCATCTGTCC	3179	
QY	3183	AGCTTGCCANT	CAACGGCCACAGTAATATGCATCAATCAGACATCTGTGAGAAGTGTCA	3242		
DB	3180	AGCTTGCCANT	CAACGGCCACAGTAATATGCATCAATCAGACATCTGTGAGAAGTGTGA	3239		
QY	3243	GAACCTGAC	CACAGGCAAGCACCTGCAGACCTGCATATCTGGCTTCTACGGTGATCCCAAC	3302		
DB	3240	GAACCTGAC	CACAGGCAAGCACCTGCAGACCTGCATATCTGGCTTCTACGGTGATCCCAAC	3299		
QY	3303	CAATGGAGGAAAT	GTGAGCCATGCAATGGGCAAGCGGTCTCTGTGTGCAACACCAA	3362		
DB	3300	CAATGGAGGAAAT	GTGAGCCATGCAATGGGCAAGCGGTCTCTGTGTGCAACACCAA	3359		
QY	3363	CACGGGCAAGT	CTGTGCAACCAAGGCGGTCAAGGGGACAGGTGCCAGTATGTGA	3422		
DB	3360	CACGGGCAAGT	CTGTGCAACCAAGGCGGTCAAGGGGACAGGTGCCAGTATGTGA	3419		
QY	3423	GGTAGAAAAAT	CGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATATACCTTCTTAT	3482		
DB	3420	GGTAGAAAAAT	CGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATATACCTTCTTAT	3479		
QY	3483	TGACTATCAGT	TTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACGCTATCAA	3542		
DB	3480	TGACTATCAGT	TTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACGCTATCAA	3539		
QY	3543	TTTTTGGCTACT	CTCAGCAAAACAAACAGGGATTTGGACATGTTCAATCAATGCCCTCCAA	3602		
DB	3540	TTTTTGGCTACT	CTCAGCAAAACAAACAGGGATTTGGACATGTTCAATCAATGCCCTCCAA	3599		
QY	3603	GAATTTCAAC	CTCAACATCACTGGGTGCCAGTTTCTCAGCTGGAAACCCAGCTGGAGA	3662		
DB	3600	GAATTTCAAC	CTCAACATCACTGGGTGCCAGTTTCTCAGCTGGAAACCCAGCTGGAGA	3659		
QY	3663	AGAGATGCTGT	TTTTCAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGA	3722		
DB	3660	AGAGATGCTGT	TTTTCAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGA	3719		
QY	3723	GAAGTTTGAT	TTTTCGCAACCAACCAATATACCTTTCTTTGTTTATGTCAAGTAATTCAC	3782		
DB	3720	GAAGTTTGAT	TTTTCGCAACCAACCAATATACCTTTCTTTGTTTATGTCAAGTAATTCAC	3779		
QY	3783	CTGGCCCATCA	AAATTCAGAT	3803		
DB	3780	CTGGCCCATCA	AAATTCAGAT	3800		

### RESULT 3

```

RES001.3
US-09-893-238-1
; Sequence 1, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21

```





QY 1787 CCTGTGACCGCTGGTCACTGCTCCAGACCTGATCTCCACCATGATGTCAACAGATTG 1846  
Db 2053 CTTGTGACCGATGTCAGTGTCTCCAGACCTGAGCTCCATCATGATGTCAACAGATTG 2112  
QY 1847 GCCATTGACAGTCTTTACACAGCACCATTGATGTGTTGGTGGTTTCAATAGTCTCC 1906  
Db 2113 GCCATTGACAGTCTTTGACACAGCACCATTGATGTGTTGGTGGTTTCAACAGCCTCC 2172  
QY 1907 TCCTCAGCGCATCTCTGGTATTTACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCG 1966  
Db 2173 TCCTCAGTGAAGTCTTGGTCTTTACCTCGGAGCAGTGGATGCACACCGCAGTGAAGCTG 2232  
QY 1967 CTTGTTTAGCAGCAGCCTGGTATTTGGTGTGTGTGGAACACAGGTCGTCTCAGTGTA 2026  
Db 2233 CTTGTGTGCGCAGCAGCCTGGTATCCGGTGTCTGTGGGACACACAGTCGTCTCGATGTA 2292  
QY 2027 TCTGTGGGCTGCGCACTGATCAACAAAGAAAGTTAAATCAGATGTTTTTCCA 2086  
Db 2293 CTTCTGGGAGTTGGCACTGGAAGAACAGCAGAAAGTTAAATCAGAGTGTCTTA 2352  
QY 2087 AAAGAACTCTTGAACATGACAGATGTGACAGCAGCAGATGTGTAAGTGTACAGCCA 2146  
Db 2353 AAAGAACTCTTGAACATGACAGATGTGACAGCAGCAGATGTGTAAGTGTACAGCCA 2412  
QY 2147 ACACCAATGACTGCCACTGGTGCATGACCATTTGTGTCGCCAGGAACACAGCTGCTCAG 2206  
Db 2413 ATACCAATGACTGCCACTGGTGCATGATCTGTGTCCTGTGGAACACAGCTGCAAG 2472  
QY 2207 AAGCCAGATCTCCATTTTAAAGTATGAGAAATGCCCCAAGGATAACCCCATGCTACT 2266  
Db 2473 AAGCCAGATCTCCATTTGCAAGTATGAGATTTGCCCAAGGATAACCCCATGCTACT 2532  
QY 2267 GTAACAGAGCAGCTGCGAGAGCTGTGCTGCGAGCAGAACTGCCAGTGGAGGCC 2326  
Db 2533 GCAATAAGAAACACAGCTGCGAGAGCTGTGCTGCGAGCAGAACTGCCAGTGGAGGCC 2592  
QY 2327 GGAATCAGAGTGCATTCCTGCTGCCGAAATATCTGTGGCATTTGGTGGCATTTGGTTG 2386  
Db 2593 GGAATCAGAGTGCATTCCTGCTGCCGAAATATCTGTGGCATTTGGTGGCATTTGGTTG 2652  
QY 2387 GAAACTCATGTTGAAATTTACTACTGCCAGGAGAAATATGACAAATGCTAAATGTTCT 2446  
Db 2653 GAAACTCATGTTGAAATTTACTACTGCCAGGAGAAATATGACAAATGCTAAATGTTCT 2712  
QY 2447 GTAGAAACCAATGCTCCCTTTGGCTCTCTTACAAACCCAGAGAGTGAATTTGTCC 2506  
Db 2713 GTAGAAACCAATGCTCCCTTTGGCTCTCTTACAAACCCAGAGAGTGAATTTGTCC 2772  
QY 2507 TTAAGCAGTGGCAATATGCAATCTCAGAGCATCTCAAGCTCACTTAAACCCCAT 2566  
Db 2773 TTAAGCAGTGGCAATATGCAATCTCAGAGCATCTCAAGCTCACTTAAACCCCAT 2832  
QY 2567 GGGTCGGCTTCGGAAGATCAATGTGCTACTGTGCTGGAGAGATGTCCTCATTTA 2626  
Db 2833 GGGTTGGTCTTCGGAAGATCAATGTGCTACTGTGCTGGAGAGATGTCCTCATTTA 2892  
QY 2627 CAAATAGTTTACTACAGTGGATGCGCTGTAGCCCAAGTGAATCTGTGGAATTT 2686  
Db 2893 CAAATAGTTTGTGCAAGTGGATGCCATCTGAGCCAGTGAATGCTGGCTCTGTGGATCT 2952  
QY 2687 TATCAGAACCCAGTACTCGGGAATGAAAGCTGCAACCTGCAATCAACCTCAATGTA 2746  
Db 2953 TGTCAAGCCTAGTACTCGGGAATTAAGGCTGCAACCTGCAATCAACCTCTCAATGCA 3012  
QY 2747 GTGTCTGTGAAGGCTGCAACCCAGTCAAGCAGTGGCGGACACCATGTGCTTGA 2806  
Db 3013 GGGTCTGTGAAGGCTGCAACCCAGTCAAGCAGTGGCGGACACCATGTGCTTGA 3072  
QY 2807 GGACAGCATGTGGAGATTSCACAGCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGA 2866  
Db 3073 GGACAGCGTGTGGAGTGCATAGCAGCAGCTCGAGTGCATGTGGTGCATACATGA 3132  
QY 2867 ACCAGTGTGCACTCCAAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTATGGAAT 2926

Db 3133 AGCAGTGTGTGACATCCCAATGCTTACGTGGCTCTCTTCCCTTTGGCCAGTATGGAAT 3192  
QY 2927 GGTATACCATGAGCACCCTGCCCTCTGAAATTTGTTTCTAGGCTACTGTACTCTAGTCAAT 2986  
Db 3193 GGTATACCATGAGCAGCTGCCCTCACTGAAATTTGTTTCTAGGCTACTGTACTCTAGC 3252  
QY 2987 GCTTGAGGAAACAGGCTGTGGCTGTACTGTATCCAGCAATACTGCAAGGAAT 3046  
Db 3253 GCTTGAGCAGCAGGCTGTGGTGTGTACTGTATCCCTAGCAATACTGGAAGGAAT 3312  
QY 3047 GCATAGAGGTTTCTTAAAGGACAGTGAAGATGCTTTCGCAAGCCCTACAGGAAT 3106  
Db 3313 GTATTGAGGCGAGCTATAAGGACCTGTGAAAGTCCCTCACAGGCTCTGAGGAATG 3372  
QY 3107 TGTATCCAAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGCAGCAGATACACTGTT 3166  
Db 3373 TGTATCCAAGCCCTTCTGAACTCCAGCATGTGTCTAGAGCAGCAGATACACTGTT 3432  
QY 3167 CTTTCAATTCACCTGCTCCAGCTTCCCAATGCAACGGCCACAGTAAATGCAATCAAGACA 3226  
Db 3433 CTTTCAATTCACCTGCTCCAGCTTCCCAATGCAACGGCCACAGTAAATGCAATCAAGACA 3492  
QY 3227 TCTGTGAGAAAGTGTGAGAACTTGACACAGGCAAGCACTGCGAGACCTTCATATCTGGCT 3286  
Db 3493 TCTGTGAGAAAGTGTGAGGACCTTGACACGGGCAAGCACTGCGAGACCTTCATATCTGGCT 3552  
QY 3287 TCTACCGTGTATCCCAATGAGGAGAAATGTCAGCATGCAAGTGAATGGCAGCGCT 3346  
Db 3553 TCTATGGTGAACCGGATTAATGAGGCAATGTGAGCATGCAAGTGAATGGCAGCGCT 3612  
QY 3347 CTCTGTGCAACACCAACAGCGGCAAGTGTCTTGTGCAACCAAGGCGCTCAAGGGGAG 3406  
Db 3613 CACTGTGCAACACCAACAGCGGCAAGTGTCTTGTGCAACCAAGGCGCTCAAGGGGAG 3672  
QY 3407 AGTGCAGCTATGTAGAGTGAATAATGCAATCAAGGAAACCTCTCAGAGGAACTGTT 3466  
Db 3673 AGTGCAGCTATGTAGAGTGAATAATGCAATCAAGGAAACCTCTCAGAGGAACTGTT 3732  
QY 3467 ATTATCTCTTCTATTGACTATCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 3526  
Db 3733 ACTATACCTCTCTCATGTACTATCACTTCACTTCACTTCACTTCACTTCACTTCACT 3792  
QY 3527 ATTACACAGCTATCAATTTTGGTGTACTCTCTGACGAAACAAACAGGAAATTTGGAATGT 3586  
Db 3793 ACTACACAGCTATCAATTTTGGTGTACTCTCTGATGAACAAACAGGAAATTTGGAATGT 3852  
QY 3587 TCATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTG 3646  
Db 3853 TCATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTG 3912  
QY 3647 GAAACCCAGGCTGGAGAGAGATGCTCTGTTTCAAAAAACCAATTAAGGAGTACAAAG 3706  
Db 3913 GAACCCAGACTGGAGAGAGTGGCTGTTTCAAAAAACCAATTAAGGAGTACAAAG 3972  
QY 3707 ATAGTTTCTCAATGAGAGTTTGTGTAATTTGGAACCAACCAATATACATTTCTTTGTTT 3766  
Db 3973 ATAGTTTCTCAATGAGAGTTTGTGTAATTTGGAACCAACCAATATACATTTCTTTGTTT 4032  
QY 3767 ATGTCAGTAATTTCACTGGCCCATCAAAATTCAGATGCTTCTCTCAGCAGCAAT 3826  
Db 4033 ATGTCAGTAATTTCACTGGCCCATCAAAATTCAGATGCTTCTCTCAGCAGCAACT 4092  
QY 3827 TTAAGCAGCTGGTACAGTCTTCTGTCATTTTCTTCTAGTGTGTTTCTCTCTCTTCTCTG 3886  
Db 4093 TCATGGACCTGGTACAGTCTTCTGTCATTTTCTTCTAGTGTGTTTCTCTCTCTCTCTG 4152  
QY 3887 TGCTGTGTGTGTTTGGAGATCAAAACAAATGTTTGGGCTCCAGACGTAGAGAGCAAC 3946  
Db 4153 TGCTGTGTGTGTTTGGAGATCAAAACAAATGTTTGGGCTCCAGACGTAGAGAGCAAC 4212  
QY 3947 TTCTTTCAGAGATGCAACAGATGCGCAGCGCTCTCTTGTGCTGTGTAATGTGCTCTGG 4006

```
Db 4213 TTCTTCGGGAGATGCAACAGATGGCCAGCGCCCTTTGCTTCTGTAAACGTTGCTTGG 4272
QY 4007 AAACAGATGAGAGGCTCTCTGATCTTATTTGGGGGAGATATAAGACTGTTCCCAACCCA 4066
Db 4273 AAACAGATGAGAGGCTCTCTGATCTTATTTGGGGGAGTATAAGACTGTTCCCAACCCA 4332
QY 4067 TTGCACTGAGCGCGTGTGTTTGGCAACAAGCGCTGCTCTCTGTGTTTGTGAGGCTCC 4126
Db 4333 TTGCACTGAGCGCGTGTGTTTGGCAACAAGCGCTGCTCTCTGTGTTTGTGAGGCTCC 4392
QY 4127 CTCGAGGCTGGTGGCATCCCTCTCTCTGGGCACTCAGGTCTTGTCTGTGGCCAGCGGCC 4186
Db 4393 CTCGAGGCTGGTGGCATCCCTCTCTCTGGGCACTCAGGTCTTGTCTGTGGCCAGCGGCC 4452
QY 4187 TGGTGGACATTTCTCAGCAGATCCGATAGTGTACAGGAGAGTCAAGGACCGGTGAA 4246
Db 4453 TGGTGGACATTTCTCAGCAGATCCGATAGTGTACAGGAGAGTCAAGGACCGGTGAA 4512
QY 4247 ACCGGAAGCAGCAGCGCCCTCTGACAGCCTGGGACCTGGCATCTGA 4290
Db 4513 ACCGGAAGCAGCAGCGCCCTCTGACAGCCTGGGACCTGGCATCTGA 4556

RESULT 4
US-09-893-238-18
; Sequence 18, Application US/09893238
; Patent No. US2002015093A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-18

Query Match 52.4%; Score 2248.2; DB 9; Length 2625;
Best Local Similarity 96.8%; Pred. No. 0;
Matches 2308; Conservative 0; Mismatches 68; Indels 9; Gaps 1;

QY 3 GGTGGCCGACGCGCGCACTGAGGCAAGGCTGAGGAGGAGGACGCGCGGACGGCAGC 62
Db 9 GGTGGCCGACGCGCGCACTGAGGCAAGGCTGAGGAGGAGGACGCGCGGACGGCAGC 68
QY 63 GCTCGGGGACGAGGCGCGGCGGCACTGGGACTGGGACTGGGACTGGGAGGCGC 122
Db 69 GCTCGGGGACGAGGCGCGGCGGCGGCACTGGGACTGGGACTGGGAGGCGC 128
QY 123 GGGGCTGGGGCGCGGCTGCGCTCCCGCGGCTGCTCTCCCGCTGGCGGCGCGGCT 182
Db 129 GGGACCGGCGCGGCTGCTCTCCCGGCGGCTGCTCTCCCGGCGGCTGCGCGGCG 188
QY 183 GCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 242
Db 189 GCTGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 239
QY 243 CGAGGCGCGGCGCGGCGGCGGCTGCTGGGCTGAGCGGCGGCGGCGGCGGCGGAG 302
Db 240 CGAGGCGCGGCGGCGGCGGCGGCTGCTGGGCTGAGCGGCGGCGGCGGCGGAG 299
```

```
QY 303 TGACCGGCCCTGTGTCAACGCGCGGTGCTGTGCAACCCCTGCGCACCGGCGCAGTGGCTGCGCC 362
Db 300 TGACCGGCCCTGTGTCAACGCGCGGTGCTGTGCAACCCCTGCGCACCGGCGCAGTGGCTGCGCC 359
QY 363 CGCGGCTGGTGGGCGAGCAATGCCAGCACTCGGGGGCGGCTTCAGACTAACTGGATC 422
Db 360 CGCGGCTGGTGGGCGAGCAATGCCAGCACTCGGGGGCGGCTTCAGACTAACTGGATC 419
QY 423 TTCTGGTGTGTGACAGATGGACCTTGAATATATAAATACAAACGAGTGCACGTGGCT 482
Db 420 TTCTGGTGTGTGACAGATGGACCTTGAATATATAAATACAAACGAGTGCACGTGGCT 479
QY 483 CATTTAAGGACGACCAATAGATATAGATCTGTTTCAATCATTTTCAATCATTTTGTACAGATG 542
Db 480 CATTTAAGGACGACCAATAGATATAGATCTGTTTCAATCATTTTCAATCATTTTGTACAGATG 539
QY 543 TAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC 602
Db 540 TAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC 599
QY 603 ATTTAGTGGCCTCATTTGTTCTGAGAGATGGCAATGAGACTGCTCCCTGAGGTTGTTGC 662
Db 600 ATTTAGTGGCCTCATTTGTTCTGAGAGATGGCAATGAGACTGCTCCCTGAGGTTGTTGC 659
QY 663 CACATCAGGTTATGCTTCTGCTCATTTTATGATGCTGCTTATATTAATTTGACTGGATT 722
Db 660 CACATCAGGTTATGCTTCTGCTCATTTTATGATGCTGCTTATATTAATTTGACTGGATT 719
QY 723 TAATATTACTTACAGTTTGTATGTTCCAAATTAATGCTCAGGCGGAGGAGTGTAA 782
Db 720 TAATATTACTTACAGTTTGTATGTTCCAAATTAATGCTCAGGCGGAGGAGTGTAA 779
QY 783 GATCAGTAATACAGCGAACTGTTGAATGTAAATGTTCTGAAAATGGAAGGTGAAGC 842
Db 780 GATCAGTAATACAGCGGATCTGTTGAATGTAAATGTTCTGAAAATGGAAGGTGAAGC 839
QY 843 ATGTGACATTTCTCACTGTACAGACAACTGTGGTTTCTCATCGAGGCACTTGCATTC 902
Db 840 ATGTGACATTTCTCACTGTACAGACAACTGTGGTTTCTCATCGAGGCACTTGCATTC 899
QY 903 AAGTGAATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAGGCTCTGAGTGTGATTC 962
Db 900 AAGTGAATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAGGCTCTGAGTGTGATTC 959
QY 963 TGTACCGACTAACAGTCAATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAG 1022
Db 960 TGTACCGACTAACAGTCAATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCCAG 1019
QY 1023 AGCATCTCATTAAGCTGTGTCATATGGAACAACTTATGTGGGTTGTGAGGATATATGTT 1082
Db 1020 AGCATCTCATTAAGCTGTGTCATATGGAACAACTTATGTGGGTTGTGAGGATATATGTT 1079
QY 1083 CAACCACTCAGATTAACATGTTCTAGCGTATGACCTTCTAGCGGAGTGGCTTCC 1142
Db 1080 CAACCACTCAGATTAACATGTTCTAGCGTATGACCTTCTAGCGGAGTGGCTTCC 1139
QY 1143 ACTAAACCGTCTGTGAAACAATGTGGTTGTGTAGATATGCTCATTTTGGCAATATACAA 1202
Db 1140 ACTAAACCGTCTGTGAAACAATGTGGTTGTGTAGATATGCTCATTTTGGCAATATACAA 1199
QY 1203 GGATAAATTTACATGTATGAGGAAAATTAATGATCACTGGAATGTGACCAATGAGTT 1262
Db 1200 GGATAAATTTACATGTATGAGGAAAATTAATGATCACTGGAATGTGACCAATGAGTT 1259
QY 1263 GAGAGTTTTTCACTTCAATTAATGAGTCAATGGTGTGTGTGACCCCTAAGGCAAGGAGCA 1322
Db 1260 GAGAGTTTTTCACTTCAATTAATGAGTCAATGGTGTGTGTGACCCCTAAGGCAAGGAGCA 1319
QY 1323 GTATCAGTGTGTGGCACTCTGCACATTTGTACACTGAGAAATGCGGAGTGGTCAT 1382
Db 1320 GTATCAGTGTGTGGCACTCTGCACATTTGTACACTGAGAAATGCGGAGTGGTCAT 1379
QY 1383 GCTGCTCATTTTGGTCACTGCTCCCTCTCTATGAGTATATTAAGCAATGTGACGGAATATGA 1442
```



QY 2954 AAAATGTTTCCAGGCTACTGCTAGTCAATGCTTGGAGCAACAGGCTGTGCTGGT 3013  
Db 721 AAAATGTTTCCAGGCTACTGCTAGTCAATGCTTGGAGCAACAGGCTGTGCTGGT 780  
QY 3014 GTACTGATCCCAAGCAATCTGCTCAAGGAAATGCTAGAGGGTTCTTATAAAGACCAAG 3073  
Db 781 GTACTGATCCCAAGCAATCTGCTCAAGGAAATGCTAGAGGGTTCTTATAAAGACCAAG 840  
QY 3074 TGAAGATGCTTCCAGGCTTCCAGGCTTCCAGGAAATCTTATCCAGGCTTCTTCAATTTCCA 3133  
Db 841 TGAAGATGCTTCCAGGCTTCCAGGCTTCCAGGAAATCTTATCCAGGCTTCTTCAATTTCCA 900  
QY 3134 GCATGTGTCTAGAGCAGCAGATCAAACTGCTCTTTCATTCACCTGTCCAGCTTGCCCAAT 3193  
Db 901 GCATGTGTCTAGAGCAGCAGATCAAACTGCTCTTTCATTCACCTGTCCAGCTTGCCCAAT 960  
QY 3194 GCAAGGCGCAAGTAAATGCTCAATCAGAGCAATCTGTGAGAAAGTGTGAGAACCTGACCA 3253  
Db 961 GCAAGGCGCAAGTAAATGCTCAATCAGAGCAATCTGTGAGAAAGTGTGAGAACCTGACCA 1020  
QY 3254 CAGGCAAGCACTGCGAGACCTGCATATCTGCTTCTACGGTGTATCCCAACCAATCGAGGGA 3313  
Db 1021 CAGGCAAGCACTGCGAGACCTGCATATCTGCTTCTACGGTGTATCCCAACCAATCGAGGGA 1080  
QY 3314 AATGTACGCAATGCAAGTGCATATGCGGACGCTCTGTGCAACCAACCAACGCGGCAAGT 3373  
Db 1081 AATGTACGCAATGCAAGTGCATATGCGGACGCTCTGTGCAACCAACCAACGCGGCAAGT 1140  
QY 3374 GCTTCTGACCAACCAAGGCTCAGGCGGAGGAGTCCAGCTATGTAGTAGAATATC 3433  
Db 1141 GCTTCTGACCAACCAAGGCTCAGGCGGAGGAGTCCAGCTATGTAGTAGAATATC 1200  
QY 3434 GATACCAAGGAAACCTCTCAGAGGAAATGTTATTTACTTCTTCTTATTTGACTATCAGT 3493  
Db 1201 GATACCAAGGAAACCTCTCAGAGGAAATGTTATTTACTTCTTCTTATTTGACTATCAGT 1260  
QY 3494 TCACCTTTAGTCTATCCAGAGAAATGATGCTATTTACAGCTATCAATTTTGTGGCTA 3553  
Db 1261 TCACCTTTAGTCTATCCAGAGAAATGATGCTATTTACAGCTATCAATTTTGTGGCTA 1320  
QY 3554 CTCCTGACCAACCAACCAAGGATTTGACATGTTTCATCAATGCTTCCAGAAATTTCAACC 3613  
Db 1321 CTCCTGACCAACCAACCAAGGATTTGACATGTTTCATCAATGCTTCCAGAAATTTCAACC 1380  
QY 3614 TCAACATCACCCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTG 3673  
Db 1381 TCAACATCACCCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTG 1440  
QY 3674 TTGTTTCAAAACCAACCAATTAAGAGTACAAAGATGTTTCTCTAATGAGAAATTTGATT 3733  
Db 1441 TTGTTTCAAAACCAACCAATTAAGAGTACAAAGATGTTTCTCTAATGAGAAATTTGATT 1500  
QY 3734 TTCGCAACCAACCAACCAATATCACTTTCTTGTATGTCAGTAATTTTCACTGCGCCCATCA 3793  
Db 1501 TTCGCAACCAACCAACCAATATCACTTTCTTGTATGTCAGTAATTTTCACTGCGCCCATCA 1560  
QY 3794 AAATTCAGATGCTTCTCAGACAGCAATTTTATGACCTGCTGATCAGTTCTTGTGA 3853  
Db 1561 AAATTCAGATGCTTCTCAGACAGCAATTTTATGACCTGCTGATCAGTTCTTGTGA 1620  
QY 3854 CTTTCTTCAGTGTGTTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3913  
Db 1621 CTTTCTTCAGTGTGTTTCT 1680  
QY 3914 AAATGTTGTGGGCTTCCAGAGTATGAGCAATTTCTTTCAGAGATGCAACAGATGGCCA 3973  
Db 1681 AAATGTTGTGGGCTTCCAGAGTATGAGCAATTTCTTTCAGAGATGCAACAGATGGCCA 1740  
QY 3974 GCCGTCCCTTGTCTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4033  
Db 1741 GCCGTCCCTTGTCTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800  
QY 4034 TTGGGGGGAGTATAAGACTGTTTCCCAACCCCAATTCAGTGGAGCCGCTGTTTGGCAACA 4093

Db 1801 TTGGGGGGAGTATAAGACTGTTTCCCAACCCCAATTCAGTGGAGCCGCTGTTTGGCAACA 1860  
QY 4094 AAGCCGCTGTCT 4153  
Db 1861 AAGCCGCTGTCT 1920  
QY 4154 CTGGGCACTCAGCT 4213  
Db 1921 CTGGGCACTCAGCT 1980  
QY 4214 TAGTGTACAAAGAGAGTCTCAGGAGCCGCTGAGAAACCGGAGAGCAGAGCCCTGACAGC 4273  
Db 1981 TAGTGTACAAAGAGAGTCTCAGGAGCCGCTGAGAAACCGGAGAGCAGAGCCCTGACAGC 2040  
QY 4274 CTGGGCACTGCACTCTGA 4290  
Db 2041 CTGGGCACTGCACTCTGA 2057  
RESULT 6  
US-09-893-238-8  
; Sequence 8, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 2419  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-893-238-8  
Query Match 35.5%; Score 1521.8; DB 9; Length 2419;  
Best Local Similarity 80.7%; Pred. No. 0;  
Matches 1895; Conservative 0; Mismatches 247; Indels 205; Gaps 2;  
QY 3 GGTGCCCGCAGCGCGGCAACTGAGGCAAGCTGAGGAGGAGGAGCGCGGCGAGCGAGC 62  
Db 82 GGTGCCCGCAGCGCGGCGCACTGAGGCGCGCTGAGGCGGAGCAGAGGACGACAGCAGC 141  
QY 63 GCTCCGCGGCGAGCGCGGCGCGCACTGAGGAGTGGAGCTGGAGCTGAGGAGGAGGAGCGC 122  
Db 142 GCTCCGCGGCGAGGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGC 201  
QY 123 GGGGCTGGGGGCGCGGCGGCGGCTGCTCCCGGCGCTGCTGCTCCACCGCTGCGGCGCAGC 182  
Db 202 GGGACCGCGCGCGCGCTGCTCTCCCGCGGCTGCTGCTGCGGGGCGCTGCTCCCGCGC 261  
QY 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242  
Db 262 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312  
QY 243 CGAGGCGCGCGCGCGCGGCGGCGGCTGCTGCGGCTCAGCGCGCAGCGCGGAGGAGATG 302  
Db 313 CGAGGCGCGCTGCGGCTGCGGCGGCGGCTGCTGCGGCTGCGGCGCGGCGGAGGAGATG 372  
QY 303 TGACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 362  
Db 373 TGACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 432

QY 363 CGCGGCTGGTGGCGAGCAATGCCAGCACTCGGGGGCGGCTTACG----- 410  
 Db 433 CACGGCTGGTGGCGAGCAATGCCAGCACTCGGGGGCGGCTTACGAGCATCTGTCTC 492  
 QY 411 ----- 410  
 Db 493 ACGCCTAATACACAGCTGTTCCGGAAGGTGAGGCTGGAGAAACAGTTCGAGGCAAGCTTC 552  
 QY 411 ----- 410  
 Db 553 GGCTACAGATAAGTTCAAGAGTAACCTGGGGCAACTGGGCTTGTCTCCAAAACCAAAA 612  
 QY 411 ----- 410  
 Db 613 TGACGAAAAGGACAAAGCTAGACTCTTTTGGGAAAATTTAGCTGACTAATTTTTCACC 672  
 QY 411 ----ACTAACTGGATCTTCTGGGTTTGTGACAGATGACCTGGAAAATTATAAATAACAAA 466  
 Db 673 GAGAACTAACTGGGCTCTTCTGGATTTGACAGATGACCTGGGAAATTATAAATAAAGA 732  
 QY 467 CGAAGTGCACTGGCTCAATGAAGGACAGGCAAAATAGAAATAGAGACTTCGTTTCAATC 526  
 Db 733 CGAAGTGCACTGGCTCAATGAAGGACAGGCAAAATAGAAATAGAGACTTCGTTTCAATC 792  
 QY 527 ATTTTGTACAGAGTGTAGTTGGGACCAATTTATATTTTATGATGGGAGCTCAATTTATG 586  
 Db 793 ATTTTGTACAGAGTGTAGTTGGGACCAATTTATATTTTATGATGGGAGCTCAATTTATG 852  
 QY 587 CACCGCTAGTTTGTGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 646  
 Db 853 CACCTGATTTGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 912  
 QY 647 TCCCTGAGGTTTGGCCACATCAGGTTATGCTTGTGCTGCAATTTTATGATGCTGCTT 706  
 Db 913 CTCCTGAGGTCACCTGCTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 972  
 QY 707 ATAAATTTCACTGGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 766  
 Db 973 ATAAATTTCACTGGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1032  
 QY 767 GCGAGGAGAGTGAAGTAACTAGTAATAGCAGGAAAATGTTGAATGGAATTTCTGAAA 826  
 Db 1033 GCGAGGAGAGTGAAGTAACTAGTAATAGCAGGAAAATGTTGAATGGAATTTCTGAAA 1092  
 QY 827 ACTGGAAGGTCAGCATGTGACATCTCTCACTGTACAGCAACTGTGGTTTCTCTATC 886  
 Db 1093 ACTGGAAGGTCAGCATGTGACATCTCTCACTGTACAGCAACTGTGGTTTCTCTATC 1152  
 QY 887 GAGGCATCTGCAATTCAGATGATGTACAGAGATGCTCTCTGCTTCTCAGACTGGCAGGTC 946  
 Db 1153 GAGGCATCTGCAATTCAGATGATGTACAGAGATGCTCTCTGCTTCTCAGACTGGCAGGTC 1212  
 QY 947 CTGGATGTTCACTTCTGTACAGCTAACAGTCAATTTTGGACTCGAGAGAAATTTCTA 1006  
 Db 1213 CTGGATGTTCACTTCTGTACAGCTAACAGTCAATTTTGGACTCGAGAGAAATTTCTG 1272  
 QY 1007 ACTTAAAGCTCCCGAGAGCATCTCAATAAGCTGTGGTCAATGGAAACATTTATGGGTTG 1066  
 Db 1273 ATTTAAAGCTCCCGAGAGCATCTCAATAAGCTGTGGTCAATGGAAACATTTATGGGTTG 1332  
 QY 1067 TTGGAGGATATATGTTCAACCACTCAGATTAATAATGGTCTACGCTATGACCTTGTCT 1126  
 Db 1333 TTGGAGGATATATGTTCAACCACTCAGATTAATAATGGTCTACGCTATGACCTTGTCT 1392  
 QY 1127 CTAGGAGTGGCTTCCACTTAAACCGTTCTGTCAACCAATGTGGTTGTTAGATATGTCATT 1186  
 Db 1393 CTAGGAGTGGCTTCCACTTAAACCGTTCTGTCAACCAATGTGGTTGTTAGATATGTCATT 1452  
 QY 1187 CTTTGGCAATTAACAGGATAAAATTTACATGATGAGGAGAAAATTTGATCCAACTGGGA 1246  
 Db 1453 CTTTGGCAATTAACAGGATAAAATTTACATGATGAGGAGAAAATTTGATCCAACTGGGA 1512

QY 1247 ATGTGACCAATGAGTTGAGAGTTTTTACATTTCAATAATGAGTCATGGGTGTTGTGACCC 1306  
 Db 1513 ACGTGACCAATGAGTTGAGAGTTTTTACATTTCAATAATGAGTCATGGGTGTTGTGACCC 1572  
 QY 1307 CTAAGGCAAGAGAGAGTATGAGTGGTGGGCACTCTGCAACATTTGACACTGAAGA 1366  
 Db 1573 CGAAAGCTAAGGATCAGTATGAGTGGTGGGCACTCTGCAACATTTGACACTGGCAT 1632  
 QY 1367 ATGGCCGAGTGGTATGCTGCTCATCTTTGGTCACTGCTCCCTCTATGATATATAAGCA 1426  
 Db 1633 CTGGCCGTTGTTGATGTTGGTTCATCTTGGGTCATTTGCCACTCTATGATATATAAGG 1692  
 QY 1427 ATGTGAGGAATATGATTTGATAAGAAACACATGAGTATATTAACACCCAGGTTGCC 1486  
 Db 1693 TTGTGCGGAATATGATTTGATAAGAAACACATGAGTATATTAACATCTCAGGTTGCTC 1752  
 QY 1487 TTGTGCAAGGGGTTACGGCCATAGCAGTGTGATGACCATAGGACCGGCTTATAGC 1546  
 Db 1753 TTGTGCAAGGGGTTATGGCCACAGTAGTGTGATGACAGGACCAAGGCTCTGTAGC 1812  
 QY 1547 TTCTGCTGGCTTACAGGCTTTTTCAGTGCATAAATGATCCGGCTTGCAGATGATCTTACC 1606  
 Db 1813 TTCTGCTGGCTTACAGGCTTTTTCAGTGCATAAATGATCCGGCTTGCAGATGATCTTACC 1872  
 QY 1607 GATATGATGATATACCCAGATGTGGAATTTCTTAAAGGACAGCCGATTTTTCGTTACT 1666  
 Db 1873 GATACGATGTGATATCTCAGATGTGGAATTTCTTAAAGGACAGCCGATTTTTCGTTACT 1932  
 QY 1667 TGCACACACTGTGATGATGAGTGGAAACCATCTGCTGTGTTGGGGGAAACACACAATG 1726  
 Db 1933 TGATACACACTGTGATGATGAGTGGAAACCATCTGCTGTGTTGGGGGAAACACACAATG 1992  
 QY 1727 ACACATCTATGAGCCATGGCCCAATGCTTCTCTCATGATTTTTCAGATTTTTCAGTCTC 1786  
 Db 1993 ACATTTCCATGAGCCACGGTGGCAATGCTTCTCTCTGACCTTTCATGGCTTATGACATG 2052  
 QY 1787 CTTGTGACCGTGGTCACTGCTTCCAGACCTGATCTCCACCATGATGTCACAGATTTG 1846  
 Db 2053 CTTGTGACCGTGGTCACTGCTTCCAGACCTGATCTCCACCATGATGTCACAGATTTG 2112  
 QY 1847 GGCATTTACAGCATTTTACACACAGCACCATGATGTTGTTGCGTGGTTTCAATAGTCTCC 1906  
 Db 2113 GGCATTTACAGCATTTTACACACAGCACCATGATGTTGTTGCGTGGTTTCAATAGTCTCC 2172  
 QY 1907 TCTCAGCGACATCTCGTATTTTACCTCGAACAGTGTGATGGGATCGAGTGAAGCCG 1966  
 Db 2173 TCTCAGTGAAGTCTTGGTCTTTTACCTCGAGCAGTGCATGACACCCAGTGAAGCTG 2232  
 QY 1967 CTTGTTTACAGCAGGACCTGGTATTTGCTGTGTTGGAACACAGGCTGCTCTCAGTGTA 2026  
 Db 2233 CTTGTTTACAGCAGGACCTGGTATTTGCTGTGTTGGAACACAGGCTGCTCTCAGTGTA 2086  
 QY 2027 TCTCGTGGGCGCTGGCAACTGATGAACAAAGAAAAGTTTAAATCAGAAATGTTTTTCCA 2086  
 Db 2293 CTTCTGGGAGTTGGCAACTGATGAACAAAGAAAAGTTTAAATCAGAAATGTTTTTCCA 2352  
 QY 2087 AAAGAACTTTGACCATGACAGATGTGACCGACACACAGATTTGTTACAGCTGTACAGCA 2146  
 Db 2353 AAAGAACTTTGACCATGACAGATGTGACCGACACACAGATTTGTTACAGCTGTACAGCA 2412  
 QY 2147 ACACCAA 2153  
 Db 2413 ATACCAA 2419

RESULT 7  
 US-10-197-824-1  
 ; Sequence 1, Application US/10197824  
 ; Publication No. US20040023219A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ORIGENE TECHNOLOGIES INC  
 ; TITLE OF INVENTION: NOVEL PROSTATE CANCER GENES  
 ; FILE REFERENCE: 9U 106 R1

/ CURRENT APPLICATION NUMBER: US/10/197,824

/ CURRENT FILING DATE: 2002-07-19

/ NUMBER OF SEQ ID NOS: 39

/ SOFTWARE: PatentIn version 3.1

/ SEQ ID NO 1

/ LENGTH: 6733

/ TYPE: DNA

/ ORGANISM: Homo sapiens

/ FEATURE:

/ NAME/KEY: CDS

/ LOCATION: (575)..(4711)

/ OTHER INFORMATION:

US-10-197-824-1

Query Match 34.1%; Score 1463.8; DB 16; Length 6733;

Best Local Similarity 62.9%; Pred. No. 0;

Matches 2490; Conservative 0; Mismatches 1377; Indels 90; Gaps 11;

QY	332	GCAACCCCTGGCAGCGCCAGTCGCTGCCCGCGCGCTGGTGGCGGAGCAATGCCAGC	391
DB	789	GCCGTGTGTCACCTCCACCTCCCTCGACCCCGCGCTGGTGGCGGAGCAATGCCAGC	848
QY	392	ACTGGGGGGCGCTTCAGACTAACTGGATCTCTCGGCTTTGTGACAGATGGACCTGGAA	451
DB	849	ACTGCCAGCGCAGGTTCAAGTTTACAGAACCTTCTGGATATTTAACAGATGCCCAATTA	908
QY	452	ATTATAAATACAAACGAGTCAGCTGGCTCATTTGAGGACAGCCAAATAGAAATATGA	511
DB	909	ACTATAAATATAAACTAAATGTACTTCGCTCAITGAAAGGCTATCCAAATCCAGTGTAA	968
QY	512	GACTTCGTTCAATCAATTTTGTACAGAGTGTAGTTGGACCAITTTATATGTTTATGATG	571
DB	969	GATTAAGATTCAATCAATTTTGTACAGATGTAGCTGGATCATATGTAUGTTTATGATG	1028
QY	572	GGGACTCAATTTATGACCGGTAGTGTGCTGAATTTAGTGGCTCATTTGTTCTCGAGAG	631
DB	1029	GAGATTCAATATATGACCTTTAATAGCTGTACTTAGTGGTTTGATAGTCCCTGAAATA	1088
QY	632	ATGGCAATGAGACTCTCCCTCAGGTTGTTGCCACATCAGGTATGCTGCTGCTGATTTT	691
DB	1089	GGGGCAATGAAACTGTGCTGAGTTGTTACTACATCTGGCTATGCACTGTACATTTT	1148
QY	692	TTAGTGATGCTGCTTAAATTTGACTGGAATTAATTTACTTACAGTTTGTATGTGTC	751
DB	1149	TTAGTGATGCTGCTGATTAATCTAACTGGTTTCAACATTTTCTAATCAATCTTTGTC	1208
QY	752	CAATAACTGCTCAGGCGGAGAGGTAGATCAGTAAATAGCAGCGAACT---GTTG	808
DB	1209	CTAACAAATGCTGCTGATGGAGGTGTAACAATAGTGTCTGTTTCCCAAGTCAAGTAT	1268
QY	809	AATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACA	868
DB	1269	ATTGTGAATGTGATAAATFACTGGAAGGTGAAGCTTGTGATATTCCTTACTGTAAAGCCA	1328
QY	869	ACTGTGTTTTCCTCATCGAGCATCTGCAATTCAGTATGTCAGAGGATGCTCCTGCT	928
DB	1329	ATTGGCGAGTCCAGATCACGGTTACTGTGACCTGACTGGAGAAAATTTATGTGCTGCA	1388
QY	929	TCTCAGACTGGCAGGCTCCTGGATGTTCAAGTCTCTGTACCAGCTTAACCACTGTTTGA	988
DB	1389	ATGATAGTTGGCAAGTCTGATTTGTTTGTGATGTTTCCCTTACTGAGTCTTACTGGA	1448
QY	989	CTCAGAGGAATATCTTAACCTAAAGC-----TCCCCAGAGCATCTCAATAAGCTGTGG	1042
DB	1449	TTCTGCCCAAGCTTAAACCCCTTCACTGCTCTGTGAGTGGGCTTCCATATAAGCAGTTT	1508
QY	1043	TCAATGGAACATTTATGTGGTGTCTGAGGATATGTTCAACCACTCAGATTAATAACA	1102
DB	1509	TACAGGGAAATTTATGTGGGTGATTTGGTATATCTTTTAACTACAGTCTTTTCAAA	1568
QY	1103	TGGTTCTAGCGTATGACTTGTCTTAGGGAGTGGCTTCCACTAAACCGTCTGTGAAACA	1162
DB	1569	TGGTCTTAAATTAACAATTTAGAAAGCAGTATATGGAATGTAGGAACTCCATCAAGGGAC	1628

QY	1163	ATGTGGTTGTTAGATATGGTCATTCTTTGGCATTATACAAGGATAAAATTTATCATGTATG	1222
DB	1629	CT---CTCCAGAGATATGGACACTCTCTTGGCTTTATATCAGGAAACATCTTTATGTATG	1685
QY	1223	GAGGAAAAATTTGATCCAA---CTGGGAATGTGACCAATGAGTTGAGTTTTCACATTCTC	1279
DB	1686	GAGGCAGAAATGAAACAAATGATGGCAATGTACAGATGAATTTATGGGTTTTTAACATAC	1745
QY	1280	ATAATGAGTCATGGGTGTTGTTGACCCCTAAAGCAAGG-----AGCAGTAGTCAG	1330
DB	1746	ATAGTCAGTCATGAGTACAAAACCTCTCTGTTCTTGACATGGTCAGCAGTATGCTG	1805
QY	1331	TGTTGGGCACTCTGCACACATTTTACACTGAGAATGCCGAGTGGTCATGCTGGTCA	1390
DB	1806	TGGAGGGACATTCAGCACATATTTATGGAGTTGCATAGATGTTGTCAATGATCATAA	1865
QY	1391	TCTTTGGTCACTGCCCTCTCTATGGATATATAGCAATGTGAGGAAATATGATTTGGATA	1450
DB	1866	TATTTGGATATTTCTGCAATATATGTTTATACAGCAGCATACAGGAATACCATATCTCAT	1925
QY	1451	AGAACACATGGAGTATATTTACACACCCAGGGTCCCTTGTGCAAGGGGTTTACGGCCATA	1510
DB	1926	CAAACACTTGGCTTGTTCAGAAACTTAAAGGAGCTATTGTACAAGGTGGATATGCCATA	1985
QY	1511	GCAGTGTTTACGACCATAGGACCGAGGCCCTATAGCTTCATGTTGGTGGCTTCAAGGCTTTCA	1570
DB	1986	CTAGTGTGTATGATGAATTAACAAAGTCCATTTATGTTTCATGGAGGGTATAAAGCATTTGC	2045
QY	1571	GTGCCAATAAGTACCGGCTTGAGATGATCTCTACCGATATGATGTGATACCCAGATGT	1630
DB	2046	CAGGGAACAAATATGATTTGGTTGTATGATCTTTTAAATATGAAGTTTAACTAGACTT	2105
QY	1631	GGACCACTTTAAGGACCGCGATTTTCCGTTACTTGCACACACTGTGATGATGAGTGTG	1690
DB	2106	GGACTATTTTGAAGAAAGTGGTGTTCAGATACCTTCATTTCAGCTGTCTTATCAATG	2165
QY	1691	GAAACATGCTGCTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCA	1750
DB	2166	GAGCTATGCTTAAATTTTGGAGAAATACCCATAATGACACTTCTCTTGAGTAAACGGTGCA	2225
QY	1751	AATGCTTCTCTTCAGATTTTCATGGGCTATGACATTTGCCTGTGACCGCTGGTCAGTGTTC	1810
DB	2226	AATGTTTTCTCCGATTTCTCGCATATGACATAGCTTGTGATGAATGGAATAACTACTAC	2285
QY	1811	CCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGACGCTTTACACACA	1870
DB	2286	CAAAACCAAAATCTTCATAGATGTCAACAGATTTGGACACTCTGCAGTAGTCAATTAACG	2345
QY	1871	GCACCATGTATGTGTTCCGTTGTTTCAATAGTCTCTCTCAGCCACATCTGTGTTATTC	1930
DB	2346	GGTCCATGTATATATTTGGGGGATTTTCTAGTGTACTCTTAAATGATATCTGTATACA	2405
QY	1931	CCTCGGAACAGTGTGATGGGCATCGAGTGAAGCGCTTGTTTAGCAGCAGACCTTGTA	1990
DB	2406	AGGCTCCAAATTTGCAAGGCTTTTACAGATGAAGAACTTTGTAAAAAATGCTGCTCCAGGA	2465
QY	1991	TTCCGTTGTGTGGAAACACAGGGTCTGCTCAGTGTATCTCTGGGCGCTGCCAACTGATG	2050
DB	2466	TAAATGTTGTTGGAAATAAATCACTGTGAAT-----CTTGGGAATCTGGGA	2513
QY	2051	AACAGAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCATCAAGAT	2110
DB	2514	ATACTAATAATATTTCTAGAGCAAAAGTCCCTCTTAAACAGCTCTTCTGATGACAGAT	2573
QY	2111	GTGACCAACACACAGATTTTACAGCTGTACAGCCCAACCAATGACTGCCACTGGTGCA	2170
DB	2574	GTTACAGATATGCAATTTGCGCAGCTGTACTGCCAATACAAATGGGTGGCCATGGGTG	2633
QY	2171	ATGACCAATGTGTCCCAAGGAACCAAGCTGCTCAGAAAGCCAGATCTCCATTTTTAGGT	2230
DB	2634	ATGCAAGAAATGCAATTTCCGGCAATAGTAACCTGAGTATGTTCTGTCAAGAACTACACCA	2693





FILE REFERENCE: 7853-237  
CURRENT APPLICATION NUMBER: US/09/893,238  
CURRENT FILING DATE: 2001-06-27  
PRIOR APPLICATION NUMBER: 09/245,041  
PRIOR FILING DATE: 1999-02-05  
PRIOR APPLICATION NUMBER: 60/093,630  
PRIOR FILING DATE: 1998-07-21  
PRIOR APPLICATION NUMBER: 60/104,978  
PRIOR FILING DATE: 1998-10-20  
NUMBER OF SEQ ID NOS: 129  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 10  
LENGTH: 1051  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-893-238-10

Query Match 14.9%; Score 640.2; DB 9; Length 1051;  
Best Local Similarity 86.1%; Pred. No. 6.2e-176;  
Matches 724; Conservative 0; Mismatches 108; Indels 9; Gaps 1;

QY 3 GGTGCGCGCAGCGCGGCAACTGAGCAAGGTGAGGAGGAGGAGCGCGGCGAGCGGCGAGC 62  
DB 216 GGTGCGCGCGCGCGCGGCGACTGAGCGCGGCTGAGGCGGAGGAGGAGCGAGCGAGCGAGC 275  
QY 63 GCTCGCGGCGAGGAGCGCGCGCGCGCACTGGGACTGGGACTGGGACTGGGACTGGGAGGCC 122  
DB 276 GCGTCGCGGCGAGGAGGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGCGAGGCC 335  
QY 123 GGGGCTGGGGCT 182  
DB 336 GGGACCG 395  
QY 183 GCTG 242  
DB 396 GCTG 446  
QY 243 CGAGCG 302  
DB 447 CGAGCG 506  
QY 303 TGACCG 362  
DB 507 TGACCG 566  
QY 363 CG 422  
DB 567 CACG 626  
QY 423 TTCTGGGTTTGTGACAGATGACCTGGGAAATTAATAATACAAACGAGGAGGAGGAGGAG 482  
DB 627 TTCTGGGTTTGTGACAGATGACCTGGGAAATTAATAATACAAACGAGGAGGAGGAGGAG 686  
QY 483 CATTGAAGGACAGCGCAAAATAGAAATAGAAATAGAAATAGAAATAGAAATAGAAATAGAA 542  
DB 687 CATTGAAGGACAGCGCAAAATAGAAATAGAAATAGAAATAGAAATAGAAATAGAAATAGAA 746  
QY 543 TAGTTGGGACCATTTATATTTATGATGGGAGCTCAATTTATGACCGGCTAGTTGCTGTC 602  
DB 747 TAGTTGGGACCATTTATATTTATGATGGGAGCTCAATTTATGACCGGCTAGTTGCTGTC 806  
QY 603 ATTTAGTGGGCTCATTTGTTCTCGAGAGAGATGGCAATGAGAGTGTCCCTCGAGGTTGTTGC 662  
DB 807 CTTTGTGGGCTCATTTGTTCTCGAGAGATGGCAATGAGAGTGTCCCTCGAGGTTGTTGC 866  
QY 663 CACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722  
DB 867 CACTTCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 926  
QY 723 TAATATTACTTACAGTTTGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 782  
DB 927 TAATATTACTTACAGTTTGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 986

QY 783 GATCAGTAATAGCAGCGAAACTGTTCAATGTGAATGTTCTGAAACTGGAAGGTGAAGC 842  
DB 987 GAGCAGTAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAACTGGAAGGGCCGG 1046  
QY 843 A 843  
DB 1047 A 1047

## RESULT 9

US-10-198-846-5869/c  
Sequence 5869, Application US/10198846  
Publication No. US20030099974A1  
GENERAL INFORMATION:  
APPLICANT: Lillie, James  
APPLICANT: Xu, Yongyao  
APPLICANT: Wang, Youzhen  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
TITLE OF INVENTION: THERAPY OF BREAST CANCER  
FILE REFERENCE: MRI-049  
CURRENT APPLICATION NUMBER: US/10/198,846  
CURRENT FILING DATE: 2002-07-18  
PRIOR APPLICATION NUMBER: 60/306,220  
PRIOR FILING DATE: 2001-07-18  
NUMBER OF SEQ ID NOS: 14084  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5869  
LENGTH: 625  
TYPE: DNA  
ORGANISM: Homo sapiens  
NAME/KEY: misc feature  
LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29  
OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-5869

Query Match 6.5%; Score 277.4; DB 14; Length 625;  
Best Local Similarity 99.6%; Pred. No. 7.1e-70;  
Matches 278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2699 GTACTCGGGGACTGAAGGCTCAACCTGCATCAACCCACTCAATGCTAGTGTCTGAAA 2758  
DB 340 GTACTCGGGGACTGAAGGCTCAACCTGCATCAACCCACTCAATGCTAGTGTCTGAAA 281  
QY 2759 GGCCTGCAAAACACAGTGTCTAAGCAGTGCAGGAGTGCAGGAGTGCCTTTGAGGACAGCATGTG 2818  
DB 280 GGCCTGCAAAACACAGTGTCTAAGCAGTGCAGGAGTGCAGGAGTGCCTTTGAGGACAGCATGTG 221  
QY 2819 GAGATTGCACAGCGGAGCTGAGTGCATCTGCTGAGCAGCAACATGAAGCAGTGTGTGG 2878  
DB 220 GAGATTGCACAGCGGAGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 161  
QY 2879 ACTCCAATGCCCTATGTGGCCCTCTTCCCTTTTGGCAGTGTATGGAATGGTATACGATGA 2938  
DB 160 ACTCCAATGCCCTATGTGGCCCTCTTCCCTTTTGGCAGTGTATGGAATGGTATACGATGA 101  
QY 2939 GACCTGCCCCCTGAAAATTTGTTCCAGGCTACTGTACCT 2977  
DB 100 GACCTGCCCCCTGAAAATTTGTTCCAGGCTACTGTACCT 62

## RESULT 10

US-09-864-761-15900  
Sequence 15900, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

**TITLE OF INVENTION:** GENE EXPRESSION ANALYSIS BY MICROARRAY

**FILE REFERENCE:** Aeomica-X-1

**CURRENT APPLICATION NUMBER:** US/09/864,761

**PRIOR FILING DATE:** 2001-05-23

**PRIOR APPLICATION NUMBER:** US 60/180,312

**PRIOR FILING DATE:** 2000-02-04

**PRIOR APPLICATION NUMBER:** US 60/207,456

**PRIOR FILING DATE:** 2000-05-26

**PRIOR APPLICATION NUMBER:** US 09/632,366

**PRIOR FILING DATE:** 2000-08-03

**PRIOR APPLICATION NUMBER:** GB 24263.6

**PRIOR FILING DATE:** 2000-10-04

**PRIOR APPLICATION NUMBER:** US 60/236,359

**PRIOR FILING DATE:** 2000-09-27

**PRIOR APPLICATION NUMBER:** PCT/US01/00666

**PRIOR FILING DATE:** 2001-01-30

**PRIOR APPLICATION NUMBER:** PCT/US01/00667

**PRIOR FILING DATE:** 2001-01-30

**PRIOR APPLICATION NUMBER:** PCT/US01/00664

**PRIOR FILING DATE:** 2001-01-30

**PRIOR APPLICATION NUMBER:** PCT/US01/00669

**PRIOR FILING DATE:** 2001-01-30

**PRIOR APPLICATION NUMBER:** PCT/US01/00665

**PRIOR FILING DATE:** 2001-01-30

**PRIOR APPLICATION NUMBER:** PCT/US01/00668

**PRIOR FILING DATE:** 2001-01-30

**PRIOR APPLICATION NUMBER:** PCT/US01/00663

**PRIOR FILING DATE:** 2001-01-30

**PRIOR APPLICATION NUMBER:** PCT/US01/00662

**PRIOR FILING DATE:** 2001-01-30

**PRIOR APPLICATION NUMBER:** PCT/US01/00661

**PRIOR FILING DATE:** 2001-01-30

**PRIOR APPLICATION NUMBER:** PCT/US01/00670

**PRIOR FILING DATE:** 2001-01-30

**PRIOR APPLICATION NUMBER:** US 60/234,687

**PRIOR FILING DATE:** 2000-09-21

**PRIOR APPLICATION NUMBER:** US 09/608,408

**PRIOR FILING DATE:** 2000-06-30

**PRIOR APPLICATION NUMBER:** US 09/774,203

**PRIOR FILING DATE:** 2001-01-29

**NUMBER OF SEQ ID NOS:** 49117

**SOFTWARE:** Annomax Sequence Listing Engine vers. 1.1

**SEQ ID NO 15900**

**LENGTH:** 531

**TYPE:** DNA

**ORGANISM:** Homo sapiens

**FEATURE:**

**OTHER INFORMATION:** MAP TO ALI32773.14

**OTHER INFORMATION:** EXPRESSED IN PLACENTA, SIGNAL = 1.3

**OTHER INFORMATION:** EXPRESSED IN LUNG, SIGNAL = 1.9

**US-09-864-761-15900**

**Query Match**      5.8%; Score 249.2; DB 9; Length 531;

**Best Local Similarity** 96.9%; Pred. No. 1.1e-61;

**Matches** 254; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

**Qy** 1195 TTATACAAGGATAAAATTTCATGTATGGAGAAGAAAATGCACCAGTGGGAATGTGCC 1254

**Dd** 141 TTTTCCTAGCAAAAAAATTTACATGTATGGAGAAGAAAATGATTAACCTGGGAATGTGCC 200

**Qy** 1255 AATGAGTTGAGAGTTTTTACCATTTCAAATAGTGATCGGGTGTTTGTGACCCCCTAAGSCA 1314

**Dd** 201 AATGAGTTGAGAGTTTTTACCATTTCAAATAGTGATCGGGTGTTTGTGACCCCCTAAGSCA 260

**Qy** 1315 AAGAGCAGTATGCGATGGTTGGGACTCTGACACACATTTTACACTGAAGANTGGCGGA 1374

**Dd** 261 AAGAGCAGTATGCGATGGTTGGGACTCTGACACACATTTTACACTGAAGANTGGCGGA 320

**Qy** 1375 GTGGTCATGCTGGTGCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAG 1434

**Dd** 321 GTGGTCATGCTGGTGCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAG 380

**Qy** 1435 GAATATGATTGGATAAGAAC 1456

**Db** 381 GAATATGATTGGTAGGTATA 402

**RESULT 11**

**US-09-736-457-966/c**

; Sequence 966, Application US/09736457

; Patent No. US20020168637A1

; GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Lodes, Michael A.

; APPLICANT: Fangter, Gary

; APPLICANT: Vedvick, Tom

; APPLICANT: Carter, Darrick

; APPLICANT: Retter, Marc

; APPLICANT: Mannion, Jane

; APPLICANT: Fan, Liqun

; APPLICANT: Wang, AiJun

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.478C15

; CURRENT APPLICATION NUMBER: US/09/736,457

; CURRENT FILING DATE: 2000-12-13

; NUMBER OF SEQ ID NOS: 1864

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 966

; LENGTH: 246

; TYPE: DNA

; ORGANISM: Homo sapien

**US-09-736-457-966**

**Query Match**      5.7%; Score 246; DB 9; Length 246;

**Best Local Similarity** 100.0%; Pred. No. 5.5e-61;

**Matches** 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Qy** 2515 CTGCGAATAATGCAGTCAATCTCACAGCATGTCCAAGTCACTTAACCCCCCATGGTCGGC 2574

**Dd** 246 CTGCGAATAATGCAGTCAATCTCACAGCATGTCCAAGTCACTTAACCCCCCATGGTCGGC 187

**Qy** 2575 CTTGCGAAGATCAATGTGTCTACTGTGTGGGAAGATATGTCCCATTTTCAAAATAGT 2634

**Dd** 186 CTTGCGAAGATCAATGTGTCTACTGTGTGGGAAGATATGTCCCATTTTCAAAATAGT 127

**Qy** 2635 TTACTACAGTGGATCCCTGTGAGCCAGTGTCTGGATTTCTGTGAATTTTATCAGAA 2694

**Dd** 126 TTACTACAGTGGATCCCTGTGAGCCAGTGTCTGGATTTCTGTGAATTTTATCAGAA 67

**Qy** 2695 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCAATCAACCCATCAATGTAGTGTCTGT 2754

**Dd** 66 CCCAGTACTCGGGGACTGAAGCTGCAACCTGCAATCAACCCATCAATGTAGTGTCTGT 7

**Qy** 2755 GAAAGG 2760

**Dd** 6 GAAAGG 1

**RESULT 12**

**US-09-902-941-966/C**

; Sequence 966, Application US/09902941

; Patent No. US20020172952A1

; GENERAL INFORMATION:

; APPLICANT: Henderson, Robert A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Marnerakis, Margarita

; APPLICANT: Carter, Darrick

; APPLICANT: Fangter, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C17

; CURRENT APPLICATION NUMBER: US/09/902,941  
; CURRENT FILING DATE: 2001-07-10

; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 966

; LENGTH: 246

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-902-941-966

Query Match 5.7%; Score 246; DB 9; Length 246;

Best Local Similarity 100.0%; Pred. No. 5.5e-61;

Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCTCTCAGAGCATGTCAGAGCTCACTTAAACCCCATGGTGGC 2574

DB 246 CTGCGAATAATGCAGTCTCTCAGAGCATGTCAGAGCTCACTTAAACCCCATGGTGGC 187

QY 2575 CTTCGGAAGATCAATGTGCTCTCTCTGCTGCTGGGAAGATATGTCCTTACAAATAGT 2634

DB 186 CTTCGGAAGATCAATGTGCTCTCTCTGCTGCTGGGAAGATATGTCCTTACAAATAGT 127

QY 2635 TTACTACAGTGCATGCTCTGAGCCAGTATGCTGATTCCTGGAATTTATCAGAA 2694

DB 126 TTACTACAGTGCATGCTCTGAGCCAGTATGCTGATTCCTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754

DB 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7

QY 2755 GAAAGG 2760

DB 6 GAAAGG 1

RESULT 13

US-09-849-626-966/c

; Sequence 966, Application US/09849626

; Publication No. US20020197669A1

; GENERAL INFORMATION:

; APPLICANT: Bangur, Chaitanya

; APPLICANT: Fanger, Gary

; APPLICANT: Wang, Aijun

; APPLICANT: Wang, Tongtong

; APPLICANT: Switzer, Anne

; APPLICANT: McNeill, Patricia

; APPLICANT: Clapper, Jonathan

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.478C16

; CURRENT APPLICATION NUMBER: US/09/849,626

; CURRENT FILING DATE: 2001-05-03

; NUMBER OF SEQ ID NOS: 1926

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 966

; LENGTH: 246

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-849-626-966

Query Match 5.7%; Score 246; DB 9; Length 246;

Best Local Similarity 100.0%; Pred. No. 5.5e-61;

Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCTCTCAGAGCATGTCAGAGCTCACTTAAACCCCATGGTGGC 2574

DB 246 CTGCGAATAATGCAGTCTCTCAGAGCATGTCAGAGCTCACTTAAACCCCATGGTGGC 187

QY 2575 CTTCGGAAGATCAATGTGCTCTCTCTGCTGCTGGGAAGATATGTCCTTACAAATAGT 2634

DB 186 CTTCGGAAGATCAATGTGCTCTCTCTGCTGCTGGGAAGATATGTCCTTACAAATAGT 127

QY 2635 TTACTACAGTGCATGCTCTGAGCCAGTATGCTGATTCCTGGAATTTATCAGAA 2694

DB 126 TTACTACAGTGCATGCTCTGAGCCAGTATGCTGATTCCTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754

DB 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7

QY 2755 GAAAGG 2760

DB 6 GAAAGG 1

DB 186 CTTCGGAAGATCAATGTGCTCTCTGCTGCTGGGAAGATATGTCCTTACAAATAGT 127

QY 2635 TTACTACAGTGCATGCTCTGAGCCAGTATGCTGGAATTTATCAGAA 2694

DB 126 TTACTACAGTGCATGCTCTGAGCCAGTATGCTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754

DB 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7

QY 2755 GAAAGG 2760

DB 6 GAAAGG 1

RESULT 14

US-10-017-754-966/c

; Sequence 966, Application US/10017754

; Publication No. US20030054363A1

; GENERAL INFORMATION:

; APPLICANT: Henderson, Robert A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Marnerakis, Margarita

; APPLICANT: Carter, Darrick

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.478C18

; CURRENT APPLICATION NUMBER: US/10/017,754

; CURRENT FILING DATE: 2001-10-29

; NUMBER OF SEQ ID NOS: 2004

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 966

; LENGTH: 246

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-017-754-966

Query Match 5.7%; Score 246; DB 14; Length 246;

Best Local Similarity 100.0%; Pred. No. 5.5e-61;

Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCTCTCAGAGCATGTCAGAGCTCACTTAAACCCCATGGTGGC 2574

DB 246 CTGCGAATAATGCAGTCTCTCAGAGCATGTCAGAGCTCACTTAAACCCCATGGTGGC 187

QY 2575 CTTCGGAAGATCAATGTGCTCTCTCTGCTGCTGGGAAGATATGTCCTTACAAATAGT 2634

DB 186 CTTCGGAAGATCAATGTGCTCTCTCTGCTGCTGGGAAGATATGTCCTTACAAATAGT 127

QY 2635 TTACTACAGTGCATGCTCTGAGCCAGTATGCTGATTCCTGGAATTTATCAGAA 2694

DB 126 TTACTACAGTGCATGCTCTGAGCCAGTATGCTGATTCCTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754

DB 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7

QY 2755 GAAAGG 2760

DB 6 GAAAGG 1

RESULT 15

US-10-872-966/c

; Sequence 966, Application US/10113872

; Publication No. US20030170255A1

GENERAL INFORMATION:  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Sleath, Paul R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C19  
; CURRENT APPLICATION NUMBER: US/10/113,872  
; CURRENT FILING DATE: 2002-03-28  
; NUMBER OF SEQ ID NOS: 2011  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 966  
; LENGTH: 246  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-113-872-966

Query Match 5.7%; Score 246; DB 14; Length 246;  
Best Local Similarity 100.0%; Pred. No. 5.5e-61;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2515 CTGCGAATATGCGAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGTGGC 2574  
Db 246 CTGCGAATATGCGAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGTGGC 187  
Qy 2575 CTTCCGAGATCAATGTCTCTACTGTCTGGGAAGATATGCCCATTTACAAATAGT 2634  
Db 186 CTTCCGAGATCAATGTCTCTACTGTCTGGGAAGATATGCCCATTTACAAATAGT 127  
Qy 2635 TTACTACAGTGGATGCCGCTCTGAGCCAGTGCATCTGGATTCTGTGAAATTTATCAGAA 2694  
Db 126 TTACTACAGTGGATGCCGCTCTGAGCCAGTGCATCTGGATTCTGTGAAATTTATCAGAA 67  
Qy 2695 CCCAGTACTCGGGACTGCAAGGCTGCACTGCATCAACCCACATGGTAGTGTCTGT 2754  
Db 66 CCCAGTACTCGGGACTGCAAGGCTGCACTGCATCAACCCACATGGTAGTGTCTGT 7  
Qy 2755 GAAAGG 2760  
Db 6 GAAAGG 1

Search completed: March 5, 2004, 13:18:46  
Job time : 962.325 secs

**This Page Blank (uspto)**



Db 300 TGACCGCGCGTGTGTCAACGGCGGTGCTGCAACCCCTGGCACCGCGCAGTGCCTGCTGCC 359  
Qy 363 CGCCGGCTGGGTGGCGAGCAATGCCAGCTGCGGGGGCGCTTCAGACTAACTGGATC 422  
Db 360 CGCCGGCTGGGTGGCGAGCAATGCCAGCTGCGGGGGCGCTTCAGACTAACTGGATC 419  
Qy 423 TTCTGGGTGTGACAGATGGACCTGGAAATTTATAAATACAAAACGAAGTGCACGTGGCT 482  
Db 420 TTCTGGGTGTGACAGATGGACCTGGAAATTTATAAATACAAAACGAAGTGCACGTGGCT 479  
Qy 483 CATTGAGGACGACCAATAGAAATAGACTGCTGGTTTCAATCATTTTGTCTACAGATG 542  
Db 480 CATTGAGGACGACCAATAGAAATAGACTGCTGGTTTCAATCATTTTGTCTACAGATG 539  
Qy 543 TAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC 602  
Db 540 TAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC 599  
Qy 603 ATTAGTGGCTCATTTGCTGAGAGAGATGGCAATGAGACTGTCCTGAGGTTGTTGC 662  
Db 600 ATTAGTGGCTCATTTGCTGAGAGAGATGGCAATGAGACTGTCCTGAGGTTGTTGC 659  
Qy 663 CACATCAGGTTATGCTTGTGCTATTTTATGATGCTGCTTATATTTGACTGGAT 722  
Db 660 CACATCAGGTTATGCTTGTGCTATTTTATGATGCTGCTTATATTTGACTGGAT 719  
Qy 723 TAAATTACTTACAGTTTATGATGCTGCTTATATTTGACTGGAT 782  
Db 720 TAAATTACTTACAGTTTATGATGCTGCTTATATTTGACTGGAT 779  
Qy 783 GATCAGTAATAGCAGCAATGCTGATGTAATGCTGCAAACTGGAAGTGAAGC 842  
Db 780 GATCAGTAATAGCAGCAATGCTGATGTAATGCTGCAAACTGGAAGTGAAGC 839  
Qy 843 ATGTGACATTCCTCACTGTACAGACAACTGTGTTTCTCATCGAGGCATCTGCAATTC 902  
Db 840 ATGTGACATTCCTCACTGTACAGACAACTGTGTTTCTCATCGAGGCATCTGCAATTC 899  
Qy 903 AAGTGATGTACAGAGATGCTCTCTCTCAGACTGCGAGGTCCTGGATGTTGAGTTCC 962  
Db 900 AAGTGATGTACAGAGATGCTCTCTCTCAGACTGCGAGGTCCTGGATGTTGAGTTCC 959  
Qy 963 TGTACCACTAACCACTGCTTGGACTCGAGAGCAATTTCTAACTTAAAGCTCCCGAG 1022  
Db 960 TGTACCACTAACCACTGCTTGGACTCGAGAGCAATTTCTAACTTAAAGCTCCCGAG 1019  
Qy 1023 AGCATCTCAAAAGCTGTGTCATGGAACATTTATGTTGGTGTGTTGGAGATATATGTT 1082  
Db 1020 AGCATCTCAAAAGCTGTGTCATGGAACATTTATGTTGGTGTGTTGGAGATATATGTT 1079  
Qy 1083 CAACCACTCAGATTATACATGTTCTAGCGTATGACCTTCTTCTAGGGAGTGGCTTCC 1142  
Db 1080 CAACCACTCAGATTATACATGTTCTAGCGTATGACCTTCTTCTAGGGAGTGGCTTCC 1139  
Qy 1143 ACTAAACCGTTCTGTGAACAACTGTGTTGATATGTTGATGTTCTTTGGCATTATACAA 1202  
Db 1140 ACTAAACCGTTCTGTGAACAACTGTGTTGATATGTTGATGTTCTTTGGCATTATACAA 1199  
Qy 1203 GGATAAATTTACATGATGAGGAAATTTGATCCACTGGAATGTGACCAATCAGTT 1262  
Db 1200 GGATAAATTTACATGATGAGGAAATTTGATTTCAACTGGAATGTGACCAATCAGTT 1259  
Qy 1263 GAGAGTTTTTCAATTCATAGTCAATGGGTGTTTGTGACCCCTTAAAGCAAGGAGCA 1322  
Db 1260 GAGAGTTTTTCAATTCATAGTCAATGGGTGTTTGTGACCCCTTAAAGCAAGGAGCA 1319  
Qy 1323 GTATGAGTGGTGGGCACTGCAACATGTTTACACTGAAGAATGGCGAGTGTGAT 1382  
Db 1320 GTATGAGTGGTGGGCACTGCAACATGTTTACACTGAAGAATGGCGAGTGTGAT 1379  
Qy 1383 GCTGGTCACTTGGTCACTGCGCTCTCTATGATATATAAGCAATGTGAGGAATATGA 1442  
Db 1380 GCTGGTCACTTGGTCACTGCGCTCTCTATGATATATAAGCAATGTGAGGAATATGA 1439

Qy 1443 TTTGGATAAGACACATGGAGTATATTACACCCAGGGTGGCTTGTGCAAGGGGTGA 1502  
Db 1440 TTTGGATAAGACACATGGAGTATATTACACCCAGGGTGGCTTGTGCAAGGGGTGA 1499  
Qy 1503 CGGCCATAGCAGTGTATTACGACCATAGGACCAAGGGCCCTATACGTTTCATGTGGCTACAA 1562  
Db 1500 CGGCCATAGCAGTGTATTACGACCATAGGACCAAGGGCCCTATACGTTTCATGTGGCTACAA 1559  
Qy 1563 GCGTTTCAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGATAC 1622  
Db 1560 GCGTTTCAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGATAC 1619  
Qy 1623 CCAGATGTGGAACCATTTTAAAGACAGCGGATTTTCCGTTACTTTCACACAGCTGTGAT 1682  
Db 1620 CCAGATGTGGAACCATTTTAAAGACAGCGGATTTTCCGTTACTTTCACACAGCTGTGAT 1679  
Qy 1683 AGTGAAGTGAACCATGCTGTTGGGGGAAACACACAAATGACATCTATGAGCCA 1742  
Db 1680 AGTGAAGTGAACCATGCTGTTGGGGGAAACACACAAATGACATCTATGAGCCA 1739  
Qy 1743 TGGCGCCAAATGCTTCTCTCAGATTTTATGSCCTATGACATTTGCTGACCGCTGTC 1802  
Db 1740 TGGCGCCAAATGCTTCTCTCAGATTTTATGSCCTATGACATTTGCTGACCGCTGTC 1799  
Qy 1803 AGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1862  
Db 1800 AGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1859  
Qy 1863 ACACACAGCACCAGTATGTTGGTGGTTCATAGTCTCTCTCCACGACATCT 1922  
Db 1860 ACACACAGCACCAGTATGTTGGTGGTTCATAGTCTCTCTCCACGACATCT 1919  
Qy 1923 GGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCGCTTGTAGCAGCAGG 1982  
Db 1920 GGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCGCTTGTAGCAGCAGG 1979  
Qy 1983 ACCTGGTATTCGGTGTGTGGAAACAGGGTGTCTCAGTGTATCTCTGGGGCGCTGGC 2042  
Db 1980 ACCTGGTATTCGGTGTGTGGAAACAGGGTGTCTCAGTGTATCTCTGGGGCGCTGGC 2039  
Qy 2043 AACTCATCAACAAAGAAAGTAAATCAGATGTTTTTCCAAAGAACTCTTGACCA 2102  
Db 2040 AACTCATCAACAAAGAAAGTAAATCAGATGTTTTTCCAAAGAACTCTTGACCA 2099  
Qy 2103 TGACAGATGTGACCAAGCACAGATTTGTTACAGCTGACAGCCCAACCAATGACTGCCA 2162  
Db 2100 TGACAGATGTGACCAAGCACAGATTTGTTACAGCTGACAGCCCAACCAATGACTGCCA 2159  
Qy 2163 CTGGTGCAATGACCAATTTGTGTCGCCAGGAACACAGCTGCTCAGAAGCCAGATCTCCAT 2222  
Db 2160 CTGGTGCAATGACCAATTTGTGTCGCCAGGAACACAGCTGCTCAGAAGCCAGATCTCCAT 2219  
Qy 2223 TTTTAGGTATGAGAAATTTGCCCCCAAGGATAACCCCATGTACTGTAAACAAAGAACCCAG 2282  
Db 2220 TTTTAGGTATGAGAAATTTGCCCCCAAGGATAACCCCATGTACTGTAAACAAAGAACCCAG 2279  
Qy 2283 CTGCAAGGCTGTGCCCTGGACCAAGACTGCCAGTGGAGCCCGGAATCAGGAGTGCAT 2342  
Db 2280 CTGCAAGGCTGTGCCCTGGACCAAGACTGCCAGTGGAGCCCGGAATCAGGAGTGCAT 2339  
Qy 2343 TGCCCTGCCCCAAAATATCTGTGGCATTTGGTTGGAACTCATGTTTGA 2402  
Db 2340 TGCCCTGCCCCAAAATATCTGTGGCATTTGGTTGGAACTCATGTTTGA 2399  
Qy 2403 AATTACTCTCCAAAGAGAAATTTATGCAATGCTTAAATTTGTTCTGTAGGAACCAATGC 2462  
Db 2400 AATTACTCTCCAAAGAGAAATTTATGCAATGCTTAAATTTGTTCTGTAGGAACCAATGC 2459  
Qy 2463 CTTTGTGCTTCTCTTACAAACCCAGAAAGGTAGAATTTGTCCTTAAGCAGCTGCCAAT 2522  
Db 2460 CTTTGTGCTTCTCTTACAAACCCAGAAAGGTAGAATTTGTCCTTAAGCAGCTGCCAAT 2519



QY 2523 AATGAGTCTCAGAGCATGCTCCAGCTCACCTTAACCCCATGGGTGGGCTTCGGAA 2582  
Db 2520 AATGAGTCTCAGAGCATGCTCCAGCTCACCTTAACCCCATGGGTGGGCTTCGGAA 2579  
QY 2583 GATCAATGTGTCTCTACTGCTGCTGGGAGATATGTCCCAATTTACAAATAGTTTACTACA 2642  
Db 2580 GATCAATGTGTCTCTACTGCTGCTGGGAGATATGTCCCAATTTACAAATAGTTTACTACA 2639  
QY 2643 GTGATCCGCTCAGAGCCAGTGTGATGCTGATGCTGGAATTTTATCAGAACCCAGTAC 2702  
Db 2640 GTGATCCGCTCAGAGCCAGTGTGATGCTGGAATTTTATCAGAACCCAGTAC 2699  
QY 2703 TCGGGAGTCAAGGCTCGCACTCTCATCAACCCCACTCAATGGTAGTGTCTGTGAAGGCC 2762  
Db 2700 TCGGGAGTCAAGGCTCGCACTCTCATCAACCCCACTCAATGGTAGTGTCTGTGAAGGCC 2759  
QY 2763 TCGAAACACAGTCTAAGCAGTCCCGGACACCATGTGCTTGGAGCAGCATGTGGAGA 2822  
Db 2760 TCGAAACACAGTCTAAGCAGTCCCGGACACCATGTGCTTGGAGCAGCATGTGGAGA 2819  
QY 2823 TTGCACAGCGGAGCTCTGAGTGCATGTGTGCGAGCAACATGAAGCAGTGTGTGACTC 2882  
Db 2820 TTGCACAGCGGAGCTCTGAGTGCATGTGTGCGAGCAACATGAAGCAGTGTGTGACTC 2879  
QY 2883 CAATGCTATGTGCT 2942  
Db 2880 CAATGCTATGTGCT 2939  
QY 2943 CTGCCCCCTGAAATTTTTCAGGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3002  
Db 2940 CTGCCCCCTGAAATTTTTCAGGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2999  
QY 3003 CTGTGGTGTGTACTGATCCAGCAATCTGCGCAAGGAAATGCTAGAGGCTTCTTA 3062  
Db 3000 CTGTGGTGTGTACTGATCCAGCAATCTGCGCAAGGAAATGCTAGAGGCTTCTTA 3059  
QY 3063 TAAAGGACAGTGAAGTGCCTTCCGAGCCCTTACAGGAAATTTCTATCCAGCCCT 3122  
Db 3060 TAAAGGACAGTGAAGTGCCTTCCGAGCCCTTACAGGAAATTTCTATCCAGCCCT 3119  
QY 3123 GCTCAATTCAGCATGTCTTAGAGGACAGAGATCAACTGCTGCTTCTTCTTCTTCTTCT 3182  
Db 3120 GCTCAATTCAGCATGTCTTAGAGGACAGAGATCAACTGCTGCTTCTTCTTCTTCTTCT 3179  
QY 3183 AGCTTGCATGCAAGCCGACAGTAATGATCAATCAGAGCATCTGTGAGAGTGTGA 3242  
Db 3180 AGCTTGCATGCAAGCCGACAGTAATGATCAATCAGAGCATCTGTGAGAGTGTGA 3239  
QY 3243 GAACCTGACACAGCAGCAGCTCGAGACCTGCATATCTGGCTTCTACGCTGATCCAC 3302  
Db 3240 GAACCTGACACAGCAGCAGCTCGAGACCTGCATATCTGGCTTCTACGCTGATCCAC 3299  
QY 3303 CAATGGAGGAAATGTCAGCCATGCAAGTGGCAAGCGCTCTCTGTGCAACACCAA 3362  
Db 3300 CAATGGAGGAAATGTCAGCCATGCAAGTGGCAAGCGCTCTCTGTGCAACACCAA 3359  
QY 3363 CAGGGCAAGTCTTCTGCAACCAAGGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3422  
Db 3360 CAGGGCAAGTCTTCTGCAACCAAGGCTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3419  
QY 3423 GGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAAATGTTATTTATATCTTCTTAT 3482  
Db 3420 GGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAAATGTTATTTATATCTTCTTAT 3479  
QY 3483 TGACTATCAGTTTCACTTTAGTCTTATCCAGGAGATGATGCTTATTTACAGCTATCAA 3542  
Db 3480 TGACTATCAGTTTCACTTTAGTCTTATCCAGGAGATGATGCTTATTTACAGCTATCAA 3539  
QY 3543 TTTTGTGCTACTCTGTCAGACAAACAGGATTTGACATGTTTCAATGCTTCCAA 3602  
Db 3540 TTTTGTGCTACTCTGTCAGACAAACAGGATTTGACATGTTTCAATGCTTCCAA 3599  
QY 3603 GAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3662

Db 3600 GAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3659  
QY 3663 AGAGATGCTGTTGTTTCAAAAACCAACATTAAAGAGTACAAAGATAGTTTCTCTAAAGA 3722  
Db 3660 AGAGATGCTGTTGTTTCAAAAACCAACATTAAAGAGTACAAAGATAGTTTCTCTAAAGA 3719  
QY 3723 GAAATTTGATTTTCCGCAACCCCAAAATATCACTTTCTTTTATGTGCTGCTGCTGCTGCT 3782  
Db 3720 GAAATTTGATTTTCCGCAACCCCAAAATATCACTTTCTTTTATGTGCTGCTGCTGCTGCT 3779  
QY 3783 CTGGCCATCAAAATCAGATGCTTCTCTCAGACAGCAATTTTATGACCTGCTGCTGCTGCT 3842  
Db 3780 CTGGCCATCAAAATCAGATGCTTCTCTCAGACAGCAATTTTATGACCTGCTGCTGCTGCT 3839  
QY 3843 GTTCTTCGTGACTTTCTTTCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3902  
Db 3840 GTTCTTCGTGACTTTCTTTCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3899  
QY 3903 GAAGATCAAAACAAAGTTGTGGGCTTCCAGAGTGAAGCACTTTCTTCGAGATGCA 3962  
Db 3900 GAAGATCAAAACAAAGTTGTGGGCTTCCAGAGTGAAGCACTTTCTTCGAGATGCA 3959  
QY 3963 ACAGATGGCCAGCGCTGCT 4022  
Db 3960 ACAGATGGCCAGCGCTGCT 4019  
QY 4023 TCTCATCTTATTTGGGGAGTATAAGACTGTTCCCAAAACCCATTCGACCTGAGCCGCTG 4082  
Db 4020 TCTCATCTTATTTGGGGAGTATAAGACTGTTCCCAAAACCCATTCGACCTGAGCCGCTG 4079  
QY 4083 TTTTGGCAACAAAGCGCTGCT 4142  
Db 4080 TTTTGGCAACAAAGCGCTGCT 4139  
QY 4143 CATCCCT 4202  
Db 4140 CATCCCT 4199  
QY 4203 GCAGATGCGATAGTGTACAAGAGAGTCAAGAGCGCTGAGAACCCGGAACGAGCAGCC 4262  
Db 4200 GCAGATGCGATAGTGTACAAGAGAGTCAAGAGCGCTGAGAACCCGGAACGAGCAGCC 4259  
QY 4263 CCTGACACAGCTGGGACCTGCATCTGA 4290  
Db 4260 CCTGACACAGCTGGGACCTGCATCTGA 4287

RESULT 2

US-09-245-041-16  
; Sequence 16 Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; FILE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 4072  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-245-041-16

Query Match

86.2%; Score 3696.2; DB 3; Length 4072;



2163 QY CTGTCGAATGACATTTGTCTCCAGGAACACAGCTGCTCAGAAAGGCCAGATCTCCAT 2222  
2160 Db CTGTCGAATGACATTTGTCTCCAGGAACACAGCTGCTCAGAAAGGCCAGATCTCCAT 2219  
2223 QY TTTTAGTATGAAATTTGCCCAAGGATTAACCCCATGTACTCTGTAAACAGAGACCAAG 2282  
2220 Db TTTTAGTATGAAATTTGCCCAAGGATTAACCCCATGTACTCTGTAAACAGAGACCAAG 2279  
2283 QY CTGAGAGAGTGTGCTGACAGAGTCTCCAGTGGAGCCCGGAATCAGAGTGCAT 2342  
2280 Db CTGAGAGAGTGTGCTGACAGAGTCTCCAGTGGAGCCCGGAATCAGAGTGCAT 2339  
2343 QY TGCCCTCCCGAAAATATCTGTGCAATTTGGCTGGCAATTTGGTTGGAAATCATGTTTGA 2402  
2340 Db TGCCCTCCCGAAAATATCTGTGCAATTTGGCTGGCAATTTGGTTGGAAATCATGTTTGA 2399  
2403 QY AATTACTACTGCCAGAGAAATATGACAAATGCTTAATTTCTGTAGGAACACCAATGC 2462  
2400 Db AATTACTACTGCCAGAGAAATATGACAAATGCTTAATTTCTGTAGGAACACCAATGC 2459  
2463 QY CCTTTTGGCTTCTCTTACAACCCAGAAAGGTAGAAATTTGTCTTTAAGCAGCTGCGAAT 2522  
2460 Db CCTTTTGGCTTCTCTTACAACCCAGAAAGGTAGAAATTTGTCTTTAAGCAGCTGCGAAT 2519  
2523 QY AATGAGTATCTCAGAGCATGTCGAAGCTCACTTAACCCCATGGTTCGGCTTCGGA 2582  
2520 Db AATGAGTATCTCAGAGCATGTCGAAGCTCACTTAACCCCATGGTTCGGCTTCGGA 2579  
2583 QY GATCAATGTGCTTACTGCTGGAGATATGCTCCCAATTTTACAATAGTTTACTACA 2642  
2580 Db GATCAATGTGCTTACTGCTGGAGATATGCTCCCAATTTTACAATAGTTTACTACA 2639  
2643 QY GTGATGCCGTCTCAGAGCATGATGCTGGAAATTTCTGTGAAATTTTATCAGAACCCAGTAC 2702  
2640 Db GTGATGCCGTCTCAGAGCATGATGCTGGAAATTTTATCAGAACCCAGTAC 2699  
2703 QY TCGGGGACTGAGAGCTGCACTGCTCAACCCACTCAATGGTAGTCTGTGAAGGCC 2762  
2700 Db TCGGGGACTGAGAGCTGCACTGCTCAACCCACTCAATGGTAGTCTGTGAAGGCC 2759  
2763 QY TGCACACACAGTGTGAAGAGTGCAGGACACCAATGTCCTTTGAGGACAGCATGTGGAGA 2822  
2760 Db TGCACACACAGTGTGAAGAGTGCAGGACACCAATGTCCTTTGAGGACAGCATGTGGAGA 2819  
2823 QY TTGCACACAGGAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGGACTC 2882  
2820 Db TTGCACACAGGAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGGACTC 2879  
2883 QY CAATGCCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGAAATGCTATACGATGAGCAC 2942  
2880 Db CAATGCCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGAAATGCTATACGATGAGCAC 2939  
2943 QY CTGCCCCCTGAAAATTTGTCAGGCTACTGTACTGCTGATGCTGATGCTGAGGACCAAGG 3002  
2940 Db CTGCCCCCTGAAAATTTGTCAGGCTACTGTACTGCTGATGCTGATGCTGAGGACCAAGG 2999  
3003 QY CTGTGGCTGTGTACTGTATCCAGCAATGCTGCAAGGGAATGCTATAGAGGTTCTCTA 3062  
3000 Db CTGTGGCTGTGTACTGTATCCAGCAATGCTGCAAGGGAATGCTATAGAGGTTCTCTA 3059  
3063 QY TAAAGACACAGTGAAGATGCTTCCAGAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3122  
3060 Db TAAAGACACAGTGAAGATGCTTCCAGAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3119  
3123 QY GCTCAATTCACAGTGTGTAGAGGACAGGATGATCACTGGTCTTTTCACTACTGCC 3182  
3120 Db GCTCAATTCACAGTGTGTAGAGGACAGGATGATCACTGGTCTTTTCACTACTGCC 3179  
3183 QY AGCTTGGCAATCAAGGCGCAGTAAATGCAATCAATCAGAGCATCTGTGAGAAAGTGTGA 3242  
3180 Db AGCTTGGCAATCAAGGCGCAGTAAATGCAATCAATCAGAGCATCTGTGAGAAAGTGTGA 3239

3243 QY GAACCTGACACAGGACGACCTGCGAGACCTGTCATATCTGGCTTCTACGGTATCCAC 3302  
3240 Db GAACCTGACACAGGACGACCTGCGAGACCTGTCATATCTGGCTTCTACGGTATCCAC 3299  
3303 QY CAATGAGGGAATTTCTGAGCCATGCAATGCAATGGGACACGCTCTCTGTGCAACACCAA 3362  
3300 Db CAATGAGGGAATTTCTGAGCCATGCAATGCAATGGGACACGCTCTCTGTGCAACACCAA 3359  
3363 QY CACGGGCAAGTGTCTTCTGACACCAAGGGGCTCAAGGGGACAGTCCAGCTATGTGA 3422  
3360 Db CACGGGCAAGTGTCTTCTGACACCAAGGGGCTCAAGGGGACAGTCCAGCTATGTGA 3419  
3423 QY GGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATCTCTCTTAT 3482  
3420 Db GGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATCTCTCTTAT 3479  
3483 QY TGAATTTCACTTCACTTTTAGTCTATCCAGGAAGATCGCTATTATCAGAGTATCAA 3542  
3480 Db TGAATTTCACTTCACTTTTAGTCTATCCAGGAAGATCGCTATTATCAGAGTATCAA 3539  
3543 QY TTTTGTGGTACTCTGAGCAACAAACAGGATTTGGACATGTTTCATCAATGCCCTCAA 3602  
3540 Db TTTTGTGGTACTCTGAGCAACAAACAGGATTTGGACATGTTTCATCAATGCCCTCAA 3599  
3603 QY GAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGAGA 3662  
3600 Db GAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGAGA 3659  
3663 QY AGAGATGCTGTTTTCAAAAACCAACATTTAAGAGTACAAAGATAGTTTCTCTAAATGA 3722  
3660 Db AGAGATGCTGTTTTCAAAAACCAACATTTAAGAGTACAAAGATAGTTTCTCTAAATGA 3719  
3723 QY GAAGTTTGAATTTTGCACACCCCAATATCACTTTCTTTTGTATGTGAGTAATTTTAC 3782  
3720 Db GAAGTTTGAATTTTGCACACCCCAATATCACTTTCTTTTGTATGTGAGTAATTTTAC 3779  
3783 QY CTGGCCCATCAAAATTCAGAT 3803  
3780 Db CTGGCCCATCAAAATTCAGAT 3800

RESULT 3  
US-09-245-041-1  
; Sequence 1, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Negle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 8827  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-1

Query Match 79.2%; Score 3398; DB 3; Length 8827;  
Best Local Similarity 86.3%; Pred. No. 0;  
Matches 3869; Conservative 0; Mismatches 410; Indels 205; Gaps 2;  
QY 3 GGTGGCGGCGCGCGCGCAACTCAGGCAAGCTGAGGAGGAGGAGCGCGCGGCGGCGGCG 62  
Db 82 GGTGGCGGCGCGCGCGCGCAACTCAGGCGCGGCTGAGGCGGAGCGGAGGAGCGGCGGCG 141



Db	2293	CTCTCTGGGAGTTGGCAAC	CTGAAGAGAACAGACGAAAGATTAA	AATCAGAGTGTGTTTTCTA	2355
Qy	2087	AAAGAACTCTTGACCAATGA	CAGATGTGACAGACACAGATTGTT	ATACGCTGTACAGCA	2146
Db	2353	AAAGAACCTTTGACCAATGA	CAGATGTGACAGACACAGATTGTT	ATACGCTGTACAGCA	2412
Qy	2147	ACACCAATGACTGCCACTGTG	CAATGTGTGTCCTCCAGGAAACAC	ACAGCTGCTCAG	2206
Db	2413	ATACCAATGACTGCCACTGTG	CAATGTGTGTCCTGTGAACCAAC	AGCTGCTCAG	2472
Qy	2207	AAGGCCAGATCTCCATTTT	TAGTATGAGAAATGCCCCAAGGA	TAAACCCCATGTA	2266
Db	2473	AAGCCAGATCTCCATTTG	CCAAATGATGAGAGTTGCCCAAG	GATAAACCCCATGTA	2532
Qy	2267	GTAAACAAGAGAACGAGCTG	CAGAGCTGTGCCCTTGACACAGAA	CTGCCAGTGGAGCCCC	2326
Db	2533	GCAATAGAAAAACGAGCTG	CAGAGCTGTGCCCTTAGACCAAA	CTGCCAGTGGAGCCCC	2592
Qy	2327	GGAAATCAGGAGTGCATTTG	CCCTGCCGAAATATCTGTGTGC	ATTTGGCTATTTGGTTG	2386
Db	2593	GGAAATCAGAGTGCATCCCT	CGCCGAAATATCTGTGTGC	ATTTGGCTATTTGGTTG	2652
Qy	2387	GAATCATGTTTGAATAATTA	CTCTGCCAAGGAGAAATTA	TATGACATGCTTAAATTTGTTCT	2446
Db	2653	GAATCATGTTTGAATAATTA	CTCTGCCAAGGAGAAATTA	TATGACATGCTTAAATTTGTTCT	2712
Qy	2447	GTAGAACCAATGCCCCTTTT	GGCTCTCTTCAACCCAGAAGAG	AGGTAGAAATTTGTTCC	2506
Db	2713	GTAGAACCAATGCCCCTTTT	GGCTCTCTTCAACCCAGAAGAG	AGGTAGAAATTTGTTCC	2772
Qy	2507	TTAAGCAGCTGCGAATAAT	CAGTCACTCAGAGCATGTCCAA	AGCTCACCTTAACCCCAT	2566
Db	2773	TTAAGCAGCTTGCATTAAT	GCATCACTCAAGATATGTCCA	AGCTCACCTTACCTCCAT	2832
Qy	2567	GGGTGCGCCTTCGAAAGAT	CNAATGTGTCTACTGGTCTG	GGAGAGATGCTCCCAATTA	2626
Db	2833	GGGTGCTTTCGAAAGATCA	NAATGTGTCTACTGGTCTG	GGAGAGATGATGTCTCCATTA	2892
Qy	2627	CAAAATAGTTTACTACAGT	GCATGCCGTCTGAGCCAGTGA	TGCTGGAATTTCTGTGGAAATTT	2686
Db	2893	CAAAATAGTTTCTCAGT	GGATGCATCTGAGCCAGTGA	TGCTGGAATTTCTGTGGAAATTT	2952
Qy	2687	TATCAGAAACCCAGTACT	CGGGAGCTGAAGGCTGCAAC	CTGCATCAACCCACTCAATGGTA	2746
Db	2953	TGTACAGACCTTAGTACT	CGGGAGTTAAAGGCTGCAAC	CTGCATCAACCCCTCTCAATGGCA	3012
Qy	2747	GTGTCTGTGAAGCCTGCA	AAACACACAGTGTGAAGCAGT	GCCGACACCACTATGTCCTTGA	2806
Db	3013	CGCTCTGTGAAGGCTGCA	AAACACACAGTGTGAAGCAGT	GCCGACACCACTATGTCCTTGA	3072
Qy	2807	GGACAGCATGTGGAGATT	GCACAGCGCAGCTCTGAGTGC	ATGTGTGTGTCAGCAACATGA	2866
Db	3073	GGACAGCGTGTGGCAGT	GCATAGCAGCTCGGAGTGCAT	GTGTGTGTGTCAGTAAACATGA	3132
Qy	2867	AGCAGTGTGGACCTCAAT	AGCCTATGTGGCTCTCTTCC	CTTTTGGCCAGTGTATGAAT	2926
Db	3133	AGCAGTGTGTGACCTCAAT	GCCTACGTGGGCTCTCTTCC	CTTTTGGCCAGTGTATGAAT	3192
Qy	2927	GGTATACCATGACACCT	GTGCCCTCGAAAAATTTGTT	CAGGCTACTGTACCTGTAGTCAAT	2986
Db	3193	GGTATACCATGACAGCT	GTGCCCTCGAAAAATTTGTT	CAGGCTACTGTACCTGTAGTCAAT	3252
Qy	2987	GCTTGGAGAACCCAGGCT	GTGGCTGTGTACTGATCC	CAGCAATACTTGGCAAGGGAAT	3046
Db	3253	GCTTGGAGCAGCCAGGCT	GTGGTGTGTACTGATCC	CAGCAATACTTGGCAAGGGAAT	3312
Qy	3047	GCATAGAGGGTTCCTATA	AAAGGACCAAGATGCCCTTC	CGCAAGCCCTTACAGGAAAT	3106
Db	3313	GTATTTAGGGCAGCTATA	AAAGGACCTGTGAAGATGCC	CTCAGGCTCTCAGGAAATG	3372
Qy	3107	TCTATCCACAGCCCTGT	CTCAATCCAGCATGTGTCT	TAGAGGACAGAGATACAACTGGT	3166
Db	3373	TGTATCCACAGCCCTGT	CTCAATCCAGCATGTGTCT	TAGAGGACAGAGATACAACTGGT	3432

QY	3167	CTTTCAATCTACTGTCAGCTTGC	CAATGC	AAAGCGG	CAAGTAAATCAATCAACAGCA	3222
DB	3433	CTTTCAATCTACTGTCAGCTTGC	CAATGC	AAAGCGG	CAAGTAAATCAATCAACAGCA	3492
QY	3227	TTCTGTGAGAAAGTGTGAGAA	CCTGAC	CAAGCAAGCACTGCGAG	AAGTGCATATCTGGCT	3286
DB	3493	TTCTGTGAGAAAGTGTGAGAA	CCTGAC	CAAGCAAGCACTGCGAG	AAGTGCATATCTGGCT	3552
QY	3287	TCTACGGTGATCCCAACCAAT	TGGAGGAAATGT	CAGCCATGCAAGTGC	CAATGGGCAAGCGT	3346
DB	3553	TCTATGTGTACCGGACTTAAT	TGGAGGCAATGT	CAGCCATGCAAGTGC	CAATGGGCAAGCGT	3612
QY	3347	CTCTGTGCAACCAACCAAC	CAAGCAAGTGC	TTCTGCAACCAAGGGCGT	CAAGGGGCAAG	3406
DB	3613	CACCTGTGCAACCAACCAAC	CAAGCAAGTGC	TTCTGCAACCAAGGGCGT	CAAGGGGCAAG	3672
QY	3407	AGTCGACGCTATGTGAGGT	TAGAAATCGAT	CAAGAAACCCCTCTC	CAGAGGAACATGTT	3466
DB	3673	AGTCCGACGCTATGTGAGGT	TAGAAATCGAT	CAAGAAACCCCTCTC	CAGAGGAACATGTT	3732
QY	3467	ATTATACTCTTTCTTAAT	GACTATCAGTT	CACCTTTAGTCTAT	CCGAGGAAGATGATCGCT	3526
DB	3733	ACTATACCTTTCTCAAT	TGACTATCAGTT	CACCTTTAGTCTAT	CCGAGGAAGATGATCGCT	3792
QY	3527	ATTACACAGCTATCAATTTT	TGTGCTACTCCT	GACGAAACAAACAGGAT	TTTCGACATGT	3586
DB	3793	ACTACACAGCTATCAATTTT	TGTGCTACTCCT	GACGAAACAAACAGGAT	TTTCGACATGT	3852
QY	3587	TCATCAATGCTCCARGAA	TTTCAACTCAACAT	TCACCTGGGTCGCAAGTT	TTCTCAGCTG	3646
DB	3853	TCATCAATGCTCCARGAA	TTTCAACTCAACAT	TCACCTGGGTCGCAAGTT	TTCTCAGCTG	3912
QY	3647	GAACCCAGGCTGGAGAGAG	ATGCTGTGTTT	CAAAAAACCAATTAAGGAGT	TACAAAG	3706
DB	3913	GAACCCAGGCTGGAGAGAG	ATGCTGTGTTT	CAAAAAACCAATTAAGGAGT	TACAAAG	3972
QY	3707	ATAGTTTCTCTPAATGAGA	AGTTGATTTT	TCGCAACCAACCAATATCACTTT	TTCTTGTTT	3766
DB	3973	ATAGTTTCTCTPAATGAGA	AGTTGATTTT	TCGCAACCAACCAATATCACTTT	TTCTTGTTT	4032
QY	3767	ATGTCAGTAATTTACACT	TGGCCCATCAAAAT	TTGAGTTGCTTCTCTCAG	CACAGCAAT	3826
DB	4033	ATGTCAGTAATTTACACT	TGGCCCATCAAAAT	TTGAGTTGCTTCTCTCAG	CACAGCAAT	4092
QY	3827	TTATGGACCTGGTACAGTT	CTTCGTGACTTT	TTTCAGTTGTTTCTCT	CTCTTTTGCTCTCTGG	3886
DB	4093	TTATGGACCTGGTACAGTT	CTTCGTGACTTT	TTTCAGTTGTTTCTCT	CTCTTTTGCTCTCTGG	4152
QY	3887	TGGCTGCTGTGGTTTGG	AGAGTCAACAAAGT	CTGCTGGGCTCCAGACG	TAGAGAGCAAC	3946
DB	4153	TGGCTGCTGTGGTTTGG	AGAGTCAACAAAGT	CTGCTGGGCTCCAGACG	TAGAGAGCAAC	4212
QY	3947	TTCTTTTCGAGAGATGCA	ACAGATGGCCGCT	CTCTTTGCTCTGTAAAT	TGTCGCTCTGG	4006
DB	4213	TTCTTTTCGAGAGATGCA	ACAGATGGCCGCT	CTCTTTGCTCTGTAAAT	TGTCGCTCTGG	4272
QY	4007	AAACAGATGAGAGGCTCT	CTCATCTTAAT	TGGGGGAGTATAAGAC	TGTTCCCAAAACCA	4066
DB	4273	AAACAGATGAGAGGCTCT	CTCATCTTAAT	TGGGGGAGTATAAGAC	TGTTCCCAAAACCA	4332
QY	4067	TTGCACTGGAGCGGTGTTT	TGGCAACAAAGCGG	TGTCCTCTCTGTGTGTTT	TGAGGCTCC	4126
DB	4333	TTGCACTGGAGCGGTGTTT	TGGCAACAAAGCGG	TGTCCTCTCTGTGTGTTT	TGAGGCTCC	4392
QY	4127	CTGAGGCTGGTGGGATCCT	CTCTCTCTCTGGG	CAGTCAAGTCTCTGCT	TGTCGCGCAGCGCC	4186
DB	4393	CTGAGGCTGGTGGGATCCT	CTCTCTCTCTGGG	CAGTCAAGTCTCTGCT	TGTCGCGCAGCGCC	4452
QY	4187	TGGTGCAATTTCTCAG	CAGATGCGGATGT	ATGATCAAGGAGAGAT	TCAGGACCGCTGAGAA	4246
DB	4453	TGGTGCAATTTCTCAG	CAGATGCGGATGT	ATGATCAAGGAGAGAT	TCAGGACCGCTGAGAA	4512

QY 4247 ACCGAGCAGCAGCCCTGACAGCTGGGACCTGCATCTGA 4290  
Db 4513 ACCGAGCAGCAGCCCTGACAGCTGGGACCTGCATCTGA 4556

## RESULT 4

US-09-245-041-18  
; Sequence 18, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 2625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-245-041-18

Query Match 52.4%; Score 2248.2; DB 3; Length 2625;

Best Local Similarity 96.8%; Pred. No. 0;

Matches 2308; Conservative 0; Mismatches 68; Indels 9; Gaps 1;

QY 3 GGTGCGCGCAGCGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGAGCGGAGC 62  
Db 9 GGTGCGCGCAGCGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGAGCGGAGC 68  
QY 63 GCTCGCGGCGAGGAGCGCGGCGGCACTGGGACTGGGACCTGGGACCTGGGAGGCGC 122  
Db 69 GCTCGCGGCGAGGAGCGCGGCGGCACTGGGACTGGGACCTGGGAGGCGC 128  
QY 123 GGGGCTGGGCGCGGCTGCGCCCTCCGCGCGCTGCTGCTCCACCGCTGGGCGCAGCGCT 182  
Db 129 GGGGCGCGCGCGGCTGCTGCTCCCGCGGCTGCTGCTGCGGCGCTGCGCGCGCGC 188  
QY 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242  
Db 189 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 239  
QY 243 CGAGGCGCGCGCGCGCGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGG 302  
Db 240 CGAGGCGCGCGCGCGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGG 299  
QY 303 TGACCGGCGCTGCTGCTCAACGCGGCTGCTGCAACCTGCGACCGCGGCGGCTGCTGCG 362  
Db 300 TGACCGGCGGCTGCTGCTCAACGCGGCTGCTGCAACCTGCGACCGCGGCGGCTGCTG 359  
QY 363 CGCGGCTGCTGCTGCGGCGGCACTGCGGCTGCGGCGGCGGCTGCTGCGGCTGCTGCG 422  
Db 360 CGCGGCTGCTGCTGCGGCGGCACTGCGGCTGCGGCGGCGGCTGCTGCGGCTGCTGCG 419  
QY 423 TTCTGGGCTTGTGACAGATGACCTGGAATATATAAATAAATAAATAAATAAATAAATA 482  
Db 420 TTCTGGGCTTGTGACAGATGACCTGGAATATATAAATAAATAAATAAATAAATAAATA 479  
QY 483 CATTTGAAGGACAGCAATAAGATAAATGAGACTGCTGCTTCAATCAATTTGCTACAGATG 542  
Db 480 CATTTGAAGGACAGCAATAAGATAAATGAGACTGCTGCTTCAATCAATTTGCTACAGATG 539  
QY 543 TAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTCTGCTG 602  
Db 540 TAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTCTGCTG 599

QY 603 ATTTAGTGGCCTCATTTGTTCTTGAGAGATGCGANTGAGCTGCTCCCTGAGGTTGTTGC 562  
Db 600 ATTTAGTGGCCTCATTTGTTCTTGAGAGATGCGANTGAGCTGCTCCCTGAGGTTGTTGC 559  
QY 663 CACATCAGGTTATGCTTGTGCTGCAATTTTATAGTGTGCTGCTTATAATTTGACTGGATT 722  
Db 660 CACATCAGGTTATGCTTGTGCTGCAATTTTATAGTGTGCTGCTTATAATTTGACTGGATT 719  
QY 723 TAATATTACTTACAGTTTGTATATGTCGAATAACTGCTCAGGCCGAGGAGGTGTA 782  
Db 720 TAATATTACTTACAGTTTGTATATGTCGAATAACTGCTCAGGCCGAGGAGGTGTA 779  
QY 783 GATCAGTAATAGCAGCAAACTGTTGAATGTGAATCTTCTGAAACTGAAAGGTGAAGC 842  
Db 780 GATCAGTAATAGCAGCAAACTGTTGAATGTGAATCTTCTGAAACTGAAAGGTGAAGC 839  
QY 843 ATGTGACATTCCTCCTACTGACAGCAAACTGTTGGTTTTCTCATGAGGCAATCTGCAATTC 902  
Db 840 ATGTGACATTCCTCCTACTGACAGCAAACTGTTGGTTTTCTCATGAGGCAATCTGCAATTC 899  
QY 903 AAGTGATGTCAGAGGATGCTGCTGCTCTCAGACTGGCAGGCTGCTGATGTTGCTGCTCC 962  
Db 900 AAGTGATGTCAGAGGATGCTGCTGCTCTCAGACTGGCAGGCTGCTGATGTTGCTGCTCC 959  
QY 963 TGTACCAGCTAACCAAGTCAATTTTGGACTGAGAGGAATATTCTAACTTTAAAGCTCCCCAG 1022  
Db 960 TGTACCAGCTAACCAAGTCAATTTTGGACTGAGAGGAATATTCTAACTTTAAAGCTCCCCAG 1019  
QY 1023 AGCATCTCATAAAGCTGCTGCTCAATGGAACAAATATGTTGGTTGTTGGAGATATATGTT 1082  
Db 1020 AGCATCTCATAAAGCTGCTGCTCAATGGAACAAATATGTTGGTTGTTGGAGATATATGTT 1079  
QY 1083 CAACCACTCAGATTAATACATGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGCTTCC 1142  
Db 1080 CAACCACTCAGATTAATACATGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGCTTCC 1139  
QY 1143 ACTAAACCGTTCTGTGAACAATGTTGTTGATATGTTGATGTTGTTGTTGTTGTTGTTGTT 1202  
Db 1140 ACTAAACCGTTCTGTGAACAATGTTGTTGATATGTTGATATGTTGTTGTTGTTGTTGTTG 1199  
QY 1203 GATATAAATTTACATGTTGAGGAGGAAATTCATCCAACTGGGATGTGACCAATGAGTT 1262  
Db 1200 GATATAAATTTACATGTTGAGGAGGAAATTCATCCAACTGGGATGTGACCAATGAGTT 1259  
QY 1263 GAGAGTTTTCATCATTAATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1322  
Db 1260 GAGAGTTTTCATCATTAATGATGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1319  
QY 1323 GTATCAGTGTGTTGGGCACTGTCACACATTTGTTACCTGAGGATGCGGAGTGGTCAAT 1382  
Db 1320 GTATCAGTGTGTTGGGCACTGTCACACATTTGTTACCTGAGGATGCGGAGTGGTCAAT 1379  
QY 1383 GCTGCTCATTTTGTGCTGCTGCTTCTATGGATATATAAGCAATGTGCGAGGAATATGA 1442  
Db 1380 GCTGCTCATTTTGTGCTGCTGCTTCTATGGATATATAAGCAATGTGCGAGGAATATGA 1439  
QY 1443 TTTGATGAAGAACAACATGAGTATATACACCCAGGCTGCTGTTGCGAGGGGTTA 1502  
Db 1440 TTTGATGAAGAACAACATGAGTATATACACCCAGGCTGCTGTTGCGAGGGGTTA 1499  
QY 1503 CGGCGCATGAGCAGTGTGTTTACGACCTAGGACCGAGGCGCTTACGTTTATGTTGCTGCTACA 1562  
Db 1500 CGGCGCATGAGCAGTGTGTTTACGACCTAGGACCGAGGCGCTTACGTTTATGTTGCTGCTACA 1559  
QY 1563 GCTTTTCTAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTTACCGATATGATGTTGATAC 1622  
Db 1560 GCTTTTCTAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTTACCGATATGATGTTGATAC 1619  
QY 1623 CCAGATGTGGAACCATTTCTTAAGGACAGCCGATTTTCCGTTTCTGTCACAGCTGTGAT 1682  
Db 1620 CCAGATGTGGAACCATTTCTTAAGGACAGCCGATTTTCCGTTTCTGTCACAGCTGTGAT 1679  
QY 1683 AGTGAGTGGAAACCATGCTGTTGTTGGGGGAAACACACAAATGACATCTATGAGCCA 1742



Db 1580 AGTGGTGGACCAATGCTGGTGGTGGGGAACACACACAAATCAGACCA 1739  
Qy 1743 TGGGCGCAATGCTGCTTCAAGATTCATGGGCTATGACATTCGCTGGTGC 1802  
Db 1740 TGGGCGCAATGCTGCTTCAAGATTCATGGGCTATGACATTCGCTGGTGC 1799  
Qy 1803 AGTGGTGGACCAATGCTGGTGGTGGGGAACACACACAAATCAGACCA 1862  
Db 1800 AGTGGTGGACCAATGCTGGTGGTGGGGAACACACACAAATCAGACCA 1859  
Qy 1863 ACACAGACCAATGCTGGTGGTGGGGAACACACACAAATCAGACCA 1922  
Db 1860 ACACAGACCAATGCTGGTGGTGGGGAACACACACAAATCAGACCA 1919  
Qy 1923 GGTATTCACCTCGGAAAGTGGTGGTGGGGAACACACACAAATCAGACCA 1982  
Db 1920 GGTATTCACCTCGGAAAGTGGTGGTGGGGAACACACACAAATCAGACCA 1979  
Qy 1983 ACCTGGTATTCGGTGGTGGTGGGGAACACACACAAATCAGACCA 2042  
Db 1980 ACCTGGTATTCGGTGGTGGTGGGGAACACACACAAATCAGACCA 2039  
Qy 2043 AACTGATGAAACAGAAAGTGGTGGTGGGGAACACACACAAATCAGACCA 2102  
Db 2040 AACTGATGAAACAGAAAGTGGTGGTGGGGAACACACACAAATCAGACCA 2099  
Qy 2103 TGACAGATGTCACAGACACAGATGTTGACAGCTGACAGCAACCAATGACCTCCA 2162  
Db 2100 TGACAGATGTCACAGACACAGATGTTGACAGCTGACAGCAACCAATGACCTCCA 2159  
Qy 2163 CTGGTGAATGACCAATGTTGCCCGAAGAACACAGCTGTCAGAGCCAGATCTCCAT 2222  
Db 2160 CTGGTGAATGACCAATGTTGCCCGAAGAACACAGCTGTCAGAGCCAGATCTCCAT 2219  
Qy 2223 TTTTAGGTATGAGATTCGCCCGAAGAACACAGCTGTCAGAGCCAGATCTCCAT 2282  
Db 2220 TTTTAGGTATGAGATTCGCCCGAAGAACACAGCTGTCAGAGCCAGATCTCCAT 2279  
Qy 2283 CTGACGAGCTGTGCTCGGACCAAGAACCTGCGAGTGGGAGCCCGGAATCAGAGTGCAT 2342  
Db 2280 CTGACGAGCTGTGCTCGGACCAAGAACCTGCGAGTGGGAGCCCGGAATCAGAGTGCAT 2339  
Qy 2343 TGGCTTGGCGGAAATATCTGCGCAATGCGCTGGCAATGCGCTGGTGG 2387  
Db 2340 TGGCTTGGCGGAGTAGGCTTGCAGGGTCACTCTGGTGGTGGG 2384

## RESULT 5

US-09-245-041-12  
; Sequence 12, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; OF OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 12  
; LENGTH: 6370  
; TYPE: DNA  
; ORGANISM: Homo sapiens

Query Match

43.2%; Score 1853; DB 3; Length 6370;

Best Local Similarity 96.3%; Pred. No. 0;  
Matches 1980; Conservative 7; Mismatches 47; Indels 23; Gaps 8;  
Qy 2257 ATGTACTACTTAAACAAGNAGACGAGCTGCAGAGCTGTGCCCTGGACAGAACTGCCAG 2316  
Db 1 ATGTACTACTTAAACAAGNAGACGAGCTGCAGAGCTGTGCCCTGGACAGAACTGCCAG 60  
Qy 2317 TGGAGCCCGGGAATCAGAGTGCATTTGCCCTGCCGAAAATATCTGTGGCAATGGCTGG 2376  
Db 61 TGGAGCCCGGGAATCAGAGTGCATTTGCCCTGCCGAAAATATCTGTGGCAATGGCTGG 120  
Qy 2377 CATTTGGTTGAAAATCTCATGTTTGAATAATCTACTGCCAGGAGAAATATGACAACTCT 2436  
Db 121 CATTTGGTTGAAAATCTCATGTTTGAATAATCTACTGCCAGGAGAAATATGACAACTCT 180  
Qy 2437 AAATTTCTCTGTAGGAACCAACAATGCCCTTTTGGCTTCTCTTACAAACCCAGAAAGSTA 2496  
Db 181 AAATTTCTCTGTAGGAACCAACAATGCCCTTTTGGCTTCTCTTACAAACCCAGAAAGSTA 240  
Qy 2497 GAATTTGCTTAAAGCAGCTGCGAATATATGAGTGCATCTCAGAGATGTCGAAGCTCAC 2556  
Db 241 GAATTTGCTTAAAGCAGCTGCGAATATATGAGTGCATCTCAGAGATGTCGAAGCTCAC 300  
Qy 2557 TTAACCCCATGGTGGGCTT---CGGAAGATCAATGTCTCTACTGG---TGCTGGG 2608  
Db 301 TTAACCCCATGGTGGGCTT---CGGAAGATCAATGTCTCTACTGG---TGCTGGG 360  
Qy 2609 AAGATATGCTCCCAATTTACAAATAG-TTTPACTACAGTGGATGCC---GTCTGAGCCAG 2663  
Db 361 GATATGCTCCCAATTTTACAAATAGTTTACTACAGTGGATGSCCGTCTTGGGCCCCAG 420  
Qy 2664 TGAATGCT---GGATTTCTGTGGAATTTT---ATCAGAACCCAGTACTCGGGAGCTGAAGG 2716  
Db 421 TGTGCTTGGATTTCTGTGGGAATTTTATTCAGGAACCCAGTTTACTTTGGGAGCTGAAGG 480  
Qy 2717 CTGCAACCTGTCAT---CAACCCCATCAATGCTAGTGTCTGTGAAGCCCTGCAAAACACAG 2774  
Db 481 CTGCAACCTGTCAT---CAACCCCATCAATGCTAGTGTCTGTGAAGCCCTGCAAAACACAG 540  
Qy 2775 TGCTAA-GCAGTGCAGGACACCATGCTTGGAGCAGCATGTGGAGATGTCACCCAGG 2833  
Db 541 TGCTAAAGGAGTGCAGGACACCATGCTTGGAGCAGCATGTGGAGATGTCACCCAGG 600  
Qy 2834 GCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGGAATCCCAATGCCATG 2893  
Db 601 GCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGGAATCCCAATGCCATG 660  
Qy 2894 TGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTTATGATGACACCTGCCCCCTG 2953  
Db 661 TGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTTATGATGACACCTGCCCCCTG 720  
Qy 2954 AAAATTTGTCAGCTACTGTACTGTAGTCAATGTTGGAGCAACCCAGGCTGTGGCTGT 3013  
Db 721 AAAATTTGTCAGCTACTGTACTGTAGTCAATGTTGGAGCAACCCAGGCTGTGGCTGT 780  
Qy 3014 GTACTCATCCAGCAATATCTGCAAGGGAATGCAATAGAGGTTCTCTATAAGAGCCAG 3073  
Db 781 GTACTCATCCAGCAATATCTGCAAGGGAATGCAATAGAGGTTCTCTATAAGAGCCAG 840  
Qy 3074 TGAAGATGCTTGCAGAGCCCTTACAGGAAATTTCTATCCAGAGCCCTGCTCAATCCA 3133  
Db 841 TGAAGATGCTTGCAGAGCCCTTACAGGAAATTTCTATCCAGAGCCCTGCTCAATCCA 900  
Qy 3134 GCATGTGTCTAGAGGACAGCAGATCAACTGTTGTTTCTTCTACTGTCCAGCTTGGCAAT 3193  
Db 901 GCATGTGTCTAGAGGACAGCAGATCAACTGTTGTTTCTTCTACTGTCCAGCTTGGCAAT 960  
Qy 3194 GCAACGGCCACAGTAAATGCATCAATCAGAGCACTGTGTAGAGTGTGAGAACCTGACCA 3253  
Db 961 GCAACGGCCACAGTAAATGCATCAATCAGAGCACTGTGTAGAGTGTGAGAACCTGACCA 1020  
Qy 3254 CAGGACGACCTGGAGACCTGCATATCTGCTTCTACGGTATCCCAATGGAGGGA 3313





```
QY 527 ATTTTGTCTACAGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGACTCAATTTATG 586
Db 793 ATTTTGTCTACAGATGTAGCTGGACCAATTTATATGTTTATGATGGGACTCAATCTACG 852
QY 587 CACCGTAGTGTCTGCAATTTAGTGGCCCTCAATGTTCTCTGAGAGAGATGCAATGAGACTG 646
Db 853 CACCTCTGATGTCTGCTTTAGTGGCCCTCAATGTTCTCTGAGAGAGATGCAATGAGACGG 912
QY 647 TCCCTGAGGTGTTGGCCACATCAGGTTATGCTTGTCTGCAATTTTATGATGCTGCTT 706
Db 913 CTCTGAGGTCACTGTCTCACTTCAGTTATGCACTGCTGCAATTTTTCAGTATGCTGCTT 972
QY 707 ATAAATTTGACTGGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 766
Db 973 ATAAATCTGACTGGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1032
QY 767 GCCGAGAGAGTGTAAAGATCAGTAATAGCAGCAATCTGTTGAATGTAATGTTCTGAAA 826
Db 1033 GCCGAGAGAGTGTAAAGAGAGTAAAGAGAGTAAAGAGAGTAAAGAGAGTAAAGAGAGT 1092
QY 827 ACTGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAATCTGTGTTTCTCTCATC 886
Db 1093 ACTGAAAGGTGAGTGTGTGACATTCCTCACTGTACAGACAATCTGTGTTTCTCTCACTC 1152
QY 887 GAGGCATCTGCAATTTCAAGTATGTACAGAGATGCTCTCTGTTCTCAGACTGCGAGGTC 946
Db 1153 GAGGCATCTGTAATGCAAGCGATACCAAGAGGTTCTCTCTGTTCTCCTCACTGCGAGGTC 1212
QY 947 CTGATGTTCACTGTTCTGTTACAGCTAACCACTCAATTTTGACTTCGAGAGCAATTTCTA 1006
Db 1213 CTGATGTTCAATTTCTGTTGCTGACCTAACCACTCAATTTTGACTTCGAGAGCAATTTCTG 1272
QY 1007 ACTTAAAGTCTCCAGAGCATCTCAATAGCTGTGCTCAATGGAACATTTATGTTGGTTG 1066
Db 1273 ATTTAAAGTCTCCAGAGCATCTCAATAGCTGTGCTCAATGGAACATTTATGTTGGTTG 1332
QY 1067 TTGGAGGATATGTTTCAACCACTCAGATTAATACAGTTCTAGCGTATGACCTGCTT 1126
Db 1333 TTGGCGGATATGTTTCAACCACTCAGATTAATACAGTTCTAGCGTATGACCTGCTT 1392
QY 1127 CTAGGAGTGGCTTCCCAATAACCGTTCTGTGAAACAATGTTGTTTATGATATGGTCATT 1186
Db 1393 CTAGGAGTGGCTTCCCAATAACCACTTCTGTGAAACAGTGTGTTGTAAGATATGGTCATT 1452
QY 1187 CTTTGGCAATTAACAGGATAAATTTACATGTATGAGGAAATTTGATCCCACTGGA 1246
Db 1453 CTTTGGCAATTAACAGGATAAATTTACATGTATGAGGAAATTTGATCCCACTGGA 1512
QY 1247 ATGTGACCAATGAGTTGAGAGTTTTCACATTCATATGAGTATGAGTGTGTTGACCC 1306
Db 1513 ACGTGACCAATGAGTTGAGAGTTTTCATATTCATATGATATGAGTGTGTTGAACTC 1572
QY 1307 CTAAAGCAAGAGAGATGATGAGTGTGGGCACTCTGCAACATTTGTTACACTGAAGA 1366
Db 1573 CGAAAGCTAAAGATCAGTATGAGTGTGGGCACTCTGCAACATTTGTTACACTGGCAT 1632
QY 1367 ATGGCCGAGTGTCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1426
Db 1633 CTGGCCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1692
QY 1427 ATGTGCAGGAATPATGTTGGATAAGAACACATGGAGTATATTAACACACCGGTTGCC 1486
Db 1693 TTGTGCAGGAATPATGTTGGATAAGAACACATGGAGTATATTAACATCTCAGGTTGCTC 1752
QY 1487 TTGTGCAGGAGTTTACGCCATACAGTGTATGACCATAGGACCAAGGCTCTATACG 1546
Db 1753 TTGTGCAGGAGTTTATGCCACAGTAGTGTATGATGACGAGCAAGGCTCTGTACG 1812
QY 1547 TTCTAGTGTGGCTACAGGCTTTTCAAGTGTGCAATTAATGAGTGTGAGTGTGAGTGTG 1606
Db 1813 TTCTAGTGTGGCTACAGGCTTTTCAAGTGTGCAATTAATGAGTGTGAGTGTGAGTGTG 1872
```

```
QY 1607 GATATGATGTGATACCCAGATGTGGACCAATCTTTAAGACAGACCGGATTTTTCGTTACT 1666
Db 1873 GATAGATGTGATGATACTCAGATGTGGACCAATCTTTAAGACAGACCGGATTTTTCGTTACT 1932
QY 1667 TCACACACAGCTGTGATAGTGTGAACCAATCTGCTGTTTGGGGGAAACACACAATG 1726
Db 1933 TGCAATACAGCTGTGATAGTGTGAACCAATCTGCTGTTTGGAGGAAACACACAATG 1992
QY 1727 ACACATCTATAGAGCCATGGCGCAATGCTTCTCTTCAAGATTTATGCGCTATGACATTG 1786
Db 1993 ACACATCTATAGAGCCATGGCGCAATGCTTCTCTTCAAGATTTATGCGCTATGACATTG 2052
QY 1787 CTTGTGACCGTGTGCTGCTTCCAGACCTGATCTCCACCATGATCTCAACAGATTTG 1846
Db 2053 CTTGTGACCGTGTGCTGCTTCCAGACCTGATCTCCACCATGATCTCAACAGATTTG 2112
QY 1847 GCCATTACAGCTCTTACACACAGACCAATGATGTGTTGTTGTTGTTTCAATGATCTCC 1906
Db 2113 GCCATTACAGCTCTTACACACAGACCAATGATGTGTTGTTGTTGTTTCAACAGCTCC 2172
QY 1907 TCTCAGCGACATCTGTTTATTCCTCGACAGTGTGTCGCGCATCGGAGTGAAGCG 1966
Db 2173 TCTCAGTACGCTTGTGTTTATTCCTCGAGAGTGTGTCGCGCATCGGAGTGAAGCG 2232
QY 1967 CTTGTTTAGCAGCAGCACTGCTGTTTTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2026
Db 2233 CTTGTTGACAGCAGCACTGCTGTTTTCGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2292
QY 2027 TCTGTTGGCGCTGGCACTGATGAAACAGAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2086
Db 2293 CTTGTTGGAGTGTGCACTGAAAGCAACAGCAAGAAAGTAAAGTAAAGTAAAGTAAAGT 2352
QY 2087 AAAGAACTCTTCACCATGACAGATGTGACCAACACACAGATTTGTTACAGCTGTACAGCCA 2146
Db 2353 AAAGAACTCTTCACCATGACAGATGTGACCAACACACAGATTTGTTACAGCTGTACAGCCA 2412
QY 2147 ACACCAA 2153
Db 2413 ATACCAA 2419
```

## RESULT 7

```
US-09-245-041-10
; Sequence 10, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Neale, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-10
```

```
Query Match 14.9%; Score 640.2; DB 3; Length 1051;
Best Local Similarity 86.1%; Pred. No. 1.4e-154;
Matches 724; Conservative 0; Mismatches 108; Indels 9; Gaps 1;
```

```
QY 3 GTTGGCGCAGCGCGCGCACTGAGGCAAGCTGAGAGAGAGCGCGCGCGCGCAGC 62
Db 216 GTTGGCGCAGCGCGCGCACTGAGGCGCGCTGAGGCGGAGCAGCAGCAGCAGC 275
```



```
Db 186 CTTCCGAAGATCAATGTCTCTACTGTGTCTGGGAAGATATGCCCCATTACAAATAGT 127
Qy 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTGTGCTGGATTCGTGGAATTTTATCAGAA 2694
Db 126 TTACTACAGTGGATCCGCTCTGAGCCAGTGTGCTGGATTCGTGGAATTTTATCAGAA 67
Qy 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
Qy 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 10
US-09-614-124B-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodges, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614.124B
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-966

Query Match 5.7%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.4e-53;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2515 CTGCGAATATGCAGTCAATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGC 2574
Db 246 CTGCGAATATGCAGTCAATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGC 187
Qy 2575 CTTCCGAAGATCAATGTCTCTACTGTGTCTGGGAAGATATGCCCCATTACAAATAGT 2634
Db 186 CTTCCGAAGATCAATGTCTCTACTGTGTCTGGGAAGATATGCCCCATTACAAATAGT 127
Qy 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTGTGCTGGATTCGTGGAATTTTATCAGAA 2694
Db 126 TTACTACAGTGGATCCGCTCTGAGCCAGTGTGCTGGATTCGTGGAATTTTATCAGAA 67
Qy 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
Qy 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 11
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
```

```
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodges, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671.325
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-966

Query Match 5.7%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.4e-53;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2515 CTGCGAATATGCAGTCAATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGC 2574
Db 246 CTGCGAATATGCAGTCAATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGC 187
Qy 2575 CTTCCGAAGATCAATGTCTCTACTGTGTCTGGGAAGATATGCCCCATTACAAATAGT 2634
Db 186 CTTCCGAAGATCAATGTCTCTACTGTGTCTGGGAAGATATGCCCCATTACAAATAGT 127
Qy 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTGTGCTGGATTCGTGGAATTTTATCAGAA 2694
Db 126 TTACTACAGTGGATCCGCTCTGAGCCAGTGTGCTGGATTCGTGGAATTTTATCAGAA 67
Qy 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
Qy 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 12
US-09-833-381-1916
; Sequence 1916, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833.381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516.448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1916
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(415)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1916

Query Match 5.0%; Score 212.8; DB 4; Length 415;
Best Local Similarity 73.6%; Pred. No. 6.5e-45;
Matches 271; Conservative 0; Mismatches 97; Indels 0; Gaps 0;
```

QY	422	CTTCTGGGTTTGTGACAGATGGA	CTTGGAAATTTATAATACAAA	CGAAGTGCACGTGC	481
Db	5	CGTCCGGATATTAA	CAGATGCCCAATTAACTATAAATATAAA	CTAAATGTACTTGC	64
QY	482	TCATTGAAGGACGACGACCAAT	AGAATAATGAGACTTCGTTTCAAT	CATTTTGTCTACAGAGT	541
Db	65	TCATTGAAGGCTATCCAA	ATGCGAGTTTAAGTTAAGATTC	AATCAATCATTTTGTCTACAGAA	124
QY	542	GTAGTTGGGACCAATTTAT	ATGTTTATGATGGGACTCAAT	TTTATGCACCGCTAGTTGCTG	601
Db	125	GTAGCTGGGATCATATG	TATGTTTATGATGGAGATTC	AATATATGACCCTTAATAGCTG	184
QY	602	CATTTAGTGGCCCTATTG	TCTCTGAGAGAGATGCCAATG	AGACTGTCCCTGAGGTTTGTG	661
Db	185	TACTTAGTGGTTTGATG	CTCCCTGAAATTAAGGGCAATG	AAACTGTGCTGAAGTTGTTA	244
QY	662	CACATACAGGTTTATGCC	TTGCTGCTATTTTATGTAGT	GCTGCTTATAATTTGACTGGAT	721
Db	245	CTACATCTGGCTATGCA	CTGTTTACATTTTTTTAGT	GATGCTGGTATAATCTAACCTGGTT	304
QY	722	TTAATATTACTTTACAG	TTTTCATATGTGTC	CAATAAATCTCTCAGCGCCGAGGAGAGTGTA	781
Db	305	TCAAATTTCTATTCAAT	CAATCTCTGTCTCTACAA	TTGCTCTGGTCAATGGGAAGTGTA	364
QY	782	AGATCAGT	789		
Db	365	CAACTAGT	372		

```

RESULT 13
US-09-245-041-5
; Sequence 5, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-5

```

Query Match	4.7%	Score 201.2;	DB 3;	Length 90050;
Best Local Similarity	81.5%;	Pred. No. 1.5e-40;		
Matches 233;	Conservative 0;	Mismatches 53;	Indels 0;	Gaps 0;
QY	1169	TTGTTAGATGTCATCTCTTTGGCAATTATACAGCATAAATTTACATGTCATGGAGGAA	1228	
Db	6931	TGGTTTTAAAAATCCTGTTTACATCTTTCTTAGGATAAAATCTACATGATCGGAGGAA	6990	
QY	1229	AAATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCATTTTCATTAATGAGT	1288	
Db	6991	AAATTGATTCACAGGGAAACGTGACCAATGAGCTGAGAGTATTTTCATTTTCATTAATGAAT	7050	
QY	1289	CATGGGTGTTGTGACCCCTAAGCGAAAGAGCAGATATGCAGTGGTGGGCACTCTGCAC	1348	
Db	7051	CATGGGTATTTGTTAACTCCGAAGCTAAGATCAGTATGCGATGGTGTGAACCTAGCAC	7110	
QY	1349	ACATGTTTACACTGAAGAAATGCCGAGTGGTCACTGCTGTCATCTTTGGTCACTGCCCTC	1408	
Db	7111	ACATGTTTACACTGGCATCTGGCCGTTGGTTCATGTTGGTCACTCTTCGGTCACTGGCCAC	7170	

```

QY      1409  TCTATGATATATAAGCAATGTGCAGGAATATGATTGGATAAGAA  1454
      |||
      7171  TCTATGATATATAAGCGTTGTGCAGGAATATGACTTGGTATGTA  7216

RESULT 14
US-09-245-041-3/C
; Sequence 3, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 17056
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-3

```

Query Watch	3.9%;	Score 169.2;	DB 3;	Length 17056;
Best Local Similarity	86.9%;	Pred. No. 9.5e-33;		
Matches 186;	Conservative 0;	Mismatches 28;	Indels 0;	Gaps 0;

  

QY	732	TTACAGTTTTGATATGTCGCAAAATAACGTCTCAGCCGAGAGAGCTGTGAAGATCAGTAA	791
Db	4090	TTGTAGTTTTGACATGTGCCGAATAATTGCTCAGCCGAGAGAGTGTGAAGCAGTAA	4031
QY	792	TAGCAGCGAAACTGTTTGAAATGTAATGTTCTGAAACTCGAAAGGTGAAGCATGTGCACAT	851
Db	4030	CAGCAGCAGCGCTGTTAGTGTGAATGTTCTGAAACTCGAAAGGGAGTCTGTCACAT	3971
QY	852	TCCTCAGCTACAGACAAGTGGTTTCCTCATCGAGGCATCTGCAATTCAGTGATCT	911
Db	3970	TCCTCAGCTACAGACAAGTGGGCTTCCTCAGGAGCATCTGTGAATGCAAGCGATAC	3911
QY	912	CAGAGATGCTCCTGTTCTCAGACTGCCAGGT	945
Db	3910	CAGAGGTGCTCCTGTTTCCTTCATCTGGCAGGT	3877

RESULT 15  
US-09-245-041-6/c  
; Sequence 6, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245, 041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093, 630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104, 978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 226  
; TYPE: DNA  
; ORGANISM: Mus musculus

Search completed: March 5, 2004, 12:20:34  
Job time : 209.649 secs

***This Page Blank (uspto)***







Db	601	GCATTTAGTGGCCATCTGTTCTCTGAGAGAGATGGCAATGAGACTGTCCTCTGAGGTTGTT	660
Qy	661	GCCACATCAGGTTATGCTTGCTGCATTTTAACTTTAGTGATGCTGCTTATAAATTTGACATGGA	720
Db	661	GCCACATCAGGTTATGCTTGCTGCATTTTAACTTTAGTGATGCTGCTTATAAATTTGACATGGA	720
Qy	721	TTTAAATATTACTTACAGTTTGTGATATGTGTCAAATAAATCTGTCAGGCCGAGAGAGTGT	780
Db	721	TTTAAATATTACTTACAGTTTGTGATATGTGTCAAATAAATCTGTCAGGCCGAGAGAGTGT	780
Qy	781	AAGATCAGTAAATACAGCGAACTGTTGAATGTCAATGTTCTGAAAACTGGAAGGTGAA	840
Db	781	AAGATCAGTAAATACAGCGAACTGTTGAATGTGTAATGTTCTGAAAACTGGAAGGTGAA	840
Qy	841	GCATGTGACATTTCTCACTGTACAGACAACTGTGGTTTCTCATCAGGCACTCTGCAAT	900
Db	841	GCATGTGACATTTCTCACTGTACAGACAACTGTGGTTTCTCATCAGGCACTCTGCAAT	900
Qy	901	TCAAGTGATGTACAGAGATGCTCTGCTTCTCAGACTGGCAGGCTCTGAGATGTTCAAGTT	960
Db	901	TCAAGTGATGTACAGAGATGCTCTGCTTCTCAGACTGGCAGGCTCTGAGATGTTCAAGTT	960
Qy	961	CCTGTACCAGCTAACCACTCATTTTGGACTCGAGAGGAATTTCTAACTTTAAAGCTCCCC	1020
Db	961	CCTGTACCAGCTAACCACTCATTTTGGACTCGAGAGGAATTTCTAACTTTAAAGCTCCCC	1020
Qy	1021	AGAGCATCTCATAAAGCTGTGGTCAATGGAAACAATTATGTGGTGTGTTGAGAGATATATG	1080
Db	1021	AGAGCATCTCATAAAGCTGTGGTCAATGGAAACAATTATGTGGTGTGTTGAGAGATATATG	1080
Qy	1081	TTCAACCACTCAGATTTAACTGGTCTAGCGTATGACCTTGCTTCTAGGAGTGGCTT	1140
Db	1081	TTCAACCACTCAGATTTAACTGGTCTAGCGTATGACCTTGCTTCTAGGAGTGGCTT	1140
Qy	1141	CCAATAAACCGTCTGTGAAACAATGTGGTGTAGATATGCTATTTCTTTGGCATATATAC	1200
Db	1141	CCAATAAACCGTCTGTGAAACAATGTGGTGTAGATATGCTATTTCTTTGGCATATATAC	1200
Qy	1201	AAGGATAAAATTTACATGTATGAGGAGAAATTCATCCAACTGGGAATGACCAATGAG	1260
Db	1201	AAGGATAAAATTTACATGTATGAGGAGAAATTCATCCAACTGGGAATGACCAATGAG	1260
Qy	1261	TTGAGAGTTTTTCACTTCAATAATGAGTCATGGGTGTGTTGACCCCTAAGGCCAAAGGAG	1320
Db	1261	TTGAGAGTTTTTCACTTCAATAATGAGTCATGGGTGTGTTGACCCCTAAGGCCAAAGGAG	1320
Qy	1321	CAGTATGCACTGGTTGGCACTCTGCACACATTTGTTTACATGGAATGCCAGTGGTC	1380
Db	1321	CAGTATGCACTGGTTGGCACTCTGCACACATTTGTTTACATGGAATGCCAGTGGTC	1380
Qy	1381	ATGCTGGTTCATCTTTGGTTCATCGCCCTCTTATCGATATATAGCAATGTGCAAGGAATAT	1440
Db	1381	ATGCTGGTTCATCTTTGGTTCATCGCCCTCTTATCGATATATAGCAATGTGCAAGGAATAT	1440
Qy	1441	GATTTGGATAAGAACACATGAGGATATATACACCCAGGGTGCCCTTGTGCAAGGGGGT	1500
Db	1441	GATTTGGATAAGAACACATGAGGATATATACACCCAGGGTGCCCTTGTGCAAGGGGGT	1500
Qy	1501	TACGGCCATAGCAGTGTGTTACGACATAGGACAGGSCCTTATACGTTTCATGTTGGCTAC	1560
Db	1501	TACGGCCATAGCAGTGTGTTACGACATAGGACAGGSCCTTATACGTTTCATGTTGGCTAC	1560
Qy	1561	AAGCTTTTCACTGTCGCAATTAAGTACCGGCTTCAGATGATCTCTACCGATATGATGGAT	1620
Db	1561	AAGCTTTTCACTGTCGCAATTAAGTACCGGCTTCAGATGATCTCTACCGATATGATGGAT	1620
Qy	1621	ACCCAGATGTGGACCATTTTAAAGACAGCGGATTTTCCGTTACTTGCACACAGCTGTG	1680
Db	1621	ACCCAGATGTGGACCATTTTAAAGACAGCGGATTTTCCGTTACTTGCACACAGCTGTG	1680
Qy	1681	ATAGTGATGGAACCATCTGCTGTTTGGGGGAAACACACAAATGACATCTTATGAGC	1740

1681	DB	ATAGT	GAGTGGAA	CCATGCTGGTGT	TTGGGGGAAACACACAA	TGACACANCTATGAGC	1740
1741	QY	CATGG	CGCAAAATGCTTCTCTT	CAGATTTTCATGGCCCTATG	CAATTCGCTGTGACCGTGG	1800	
1741	DB	CATGG	CGCAAAATGCTTCTCTT	CAGATTTTCATGGCCCTATG	CAATTCGCTGTGACCGTGG	1800	
1801	QY	TCAGT	GCTTCCAGACCTGATCTC	CAACATGATGTCAACAGATTTGGCCATTCAGCAGTC	1860		
1801	DB	TCAGT	GCTTCCAGACCTGATCTC	CAACATGATGTCAACAGATTTGGCCATTCAGCAGTC	1860		
1861	QY	TTACA	CAACAGCACATGATGTGTT	CGGTGGTTCATATAGTCTCCCTC	CAGCGAATC	1920	
1861	DB	TTACA	CAACAGCACATGATGTGTT	CGGTGGTTCATATAGTCTCCCTC	CAGCGAATC	1920	
1921	QY	CTGGT	ATTTCACCTCGAACAGTGTG	ATGCGCATCGAGTGAAGCGCGTGTGTTT	TAGCAGCA	1980	
1921	DB	CTGGT	ATTTCACCTCGAACAGTGTG	ATGCGCATCGAGTGAAGCGCGTGTGTTT	TAGCAGCA	1980	
1981	QY	GGACT	GTGATTCGGTGTGTGTG	GAAACACAGGGTCTCTCAGTCTATCT	CGTGGCGGCTG	2040	
1981	DB	GGACT	GTGATTCGGTGTGTGTG	GAAACACAGGGTCTCTCAGTCTATCT	CGTGGCGGCTG	2040	
2041	QY	GCAAT	CGATGAACCAAGAGAAAGT	TAAAAATCAGAAATGTTTTT	TCCAAAAAGAACTCTTGAC	2100	
2041	DB	GCAAT	CGATGAACCAAGAGAAAGT	TAAAAATCAGAAATGTTTTT	TCCAAAAAGAACTCTTGAC	2100	
2101	QY	CATC	ACAGATGTGACACGACACACAGATTTGTTT	TACAGCTGTACAGCCAAACCAAA	TGACTGC	2160	
2101	DB	CATC	ACAGATGTGACACGACACACAGATTTGTTT	TACAGCTGTACAGCCAAACCAAA	TGACTGC	2160	
2161	QY	CAC	TGTGCAATGATCACCATTG	TGTCCTCCAGGAAACACACAGCTGCT	CAGAAAGGCAGATCTCC	2220	
2161	DB	CAC	TGTGCAATGATCACCATTG	TGTCCTCCAGGAAACACACAGCTGCT	CAGAAAGGCAGATCTCC	2220	
2221	QY	ATTTTT	TAGGTATGAGATTTGCC	CCCAAGATACCCCATGTACTACTGT	TACAAAGAGACC	2280	
2221	DB	ATTTTT	TAGGTATGAGATTTGCC	CCCAAGATACCCCATGTACTACTGT	TACAAAGAGACC	2280	
2281	QY	AGCT	GCAGGAGCTGTGCC	TGCGACAGAACTGCCAGTGGGAGCCCCGGAA	TACAGAGTGC	2340	
2281	DB	AGCT	GCAGGAGCTGTGCC	TGCGACAGAACTGCCAGTGGGAGCCCCGGAA	TACAGAGTGC	2340	
2341	QY	ATTG	CCCTGCCGAAAAATATCTG	TGGCAATTTGGTGGCAATTCATGTTT	G	2400	
2341	DB	ATTG	CCCTGCCGAAAAATATCTG	TGGCAATTTGGTGGCAATTCATGTTT	G	2400	
2401	QY	AAAAT	TACTACTGCCAACAGGAGAA	TATGACAACTCTAAATTTGTTCTGT	TAGGAAACCAAT	2460	
2401	DB	AAAAT	TACTACTGCCAACAGGAGAA	TATGACAACTCTAAATTTGTTCTGT	TAGGAAACCAAT	2460	
2461	QY	GCCCT	TTTTGGCTTCTTCTTAC	AACCCAGAAAGGTAGAAATTTGCTT	TAAAGCAGCTCGCA	2520	
2461	DB	GCCCT	TTTTGGCTTCTTCTTAC	AACCCAGAAAGGTAGAAATTTGCTT	TAAAGCAGCTCGCA	2520	
2521	QY	ATAAT	GCAGTCAATCTCAGACATGT	CAAGCTCACTTAACCCCATGGTGGCGCTTCGG	2580		
2521	DB	ATAAT	GCAGTCAATCTCAGACATGT	CAAGCTCACTTAACCCCATGGTGGCGCTTCGG	2580		
2581	QY	AAGAT	CAATGTGCTACTGTG	TGCGAAGATATGTCCCCATTTTACAAAT	TAGTTTACTA	2640	
2581	DB	AAGAT	CAATGTGCTACTGTG	TGCGAAGATATGTCCCCATTTTACAAAT	TAGTTTACTA	2640	
2641	QY	CAGT	GGATGCCGTGTGACCC	CAGTGTATCTGTGGAAATTTTAT	TACAAACCCAGT	2700	
2641	DB	CAGT	GGATGCCGTGTGACCC	CAGTGTATCTGTGGAAATTTTAT	TACAAACCCAGT	2700	
2701	QY	ACT	CGGGACTGAAGGCTG	CAACTGCATCAACCCCATCAAT	TGTTAGTGTCTGTGAAAGG	2760	
2701	DB	ACT	CGGGACTGAAGGCTG	CAACTGCATCAACCCCATCAAT	TGTTAGTGTCTGTGAAAGG	2760	
2761	QY	CCTG	CAAAACCACTGCTTAAG	CACTGCCGACACCAATGTG	CGCTTTGAGCAGCATGTGGA	2820	
2761	DB	CCTG	CAAAACCACTGCTTAAG	CACTGCCGACACCAATGTG	CGCTTTGAGCAGCATGTGGA	2820	

DEFINITION		Sequence 14 from patent US 6274339.	
ACCESSION		AR164816	
VERSION		AR164816.1 GI:16238025	
KEYWORDS		Unknown.	
SOURCE		Unknown.	
ORGANISM		Unknown.	
REFERENCE		1 (bases 1 to 8589)	
AUTHORS		Moore, K. and Nagle, D. Lynn.	
TITLE		Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity	
JOURNAL		Patent: US 6274339-A 14 AUG-2001;	
FEATURES		Location/Qualifiers	
source		1..8589	
		/organism="unknown"	
		/mol_type="unassigned DNA"	
ORIGIN		Query Match 88.0%; Score 3774; DB 6; Length 8589;	
		Best Local Similarity 99.9%; Pred. No. 0;	
		Matches 3974; Conservative 0; Mismatches 4; Indels 0; Gaps 0;	
QY	313	TGTGTCAACGGCGGTGCTGCAACCCCTGGCACCCGCGCGCTGCTGCGCCCGCGGTGG	372
DB	310	TGTGTCAACGGCGGTGCTGCAACCCCTGGCACCCGCGCGCTGCTGCGCCCGCGGTGG	369
QY	373	GTGGCGGAGCAATGCGACACTGCGGGGGCGCTTCAGACTAACTGGATCTCTGGGTTT	432
DB	370	GTGGCGGAGCAATGCGACACTGCGGGGGCGCTTCAGACTAACTGGATCTCTGGGTTT	429
QY	433	GTGACAGATGACCTCGGAAATTTATAATACAAACGAAAGTCAGCTGCTCAATTGAAGGA	492
DB	430	GTGACAGATGACCTCGGAAATTTATAATACAAACGAAAGTCAGCTGCTCAATTGAAGGA	489
QY	493	CAGCAAAATAGAAATAGAGACTTCGTTTCAATCAATTTTGTCTACAGATGAGTGGAC	552
DB	490	CAGCAAAATAGAAATAGAGACTTCGTTTCAATCAATTTTGTCTACAGATGAGTGGAC	549
QY	553	CATTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGC	612
DB	550	CATTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGC	609
QY	613	CTCATTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT	672
DB	610	CTCATTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT	669
QY	673	TATGCTTGTGCTGCAATTTTATGATGCTGCTTATATTTGACCTGGATTTAATATTACT	732
DB	670	TATGCTTGTGCTGCAATTTTATGATGCTGCTTATATTTGACCTGGATTTAATATTACT	729
QY	733	TACAGTTTGTATGATGTCCTCAAAATACTGCTCAGCGCGAGGAGAGTGAAGATCAGTAAT	792
DB	730	TACAGTTTGTATGATGTCCTCAAAATACTGCTCAGCGCGAGGAGAGTGAAGATCAGTAAT	789
QY	793	AGCAGGAAACTGTTGTAATGTAATGTTCTGAAACTGGAAGGTGAGCATGTGACATT	852
DB	790	AGCAGGAACTGTTGTAATGTAATGTTCTGAAACTGGAAGGTGAGCATGTGACATT	849
QY	853	CCTCACTGTACAGACAACTGTGTTTCTCTCATCGAGCATCTGCAATTCAGATGATGTC	912
DB	850	CCTCACTGTACAGACAACTGTGTTTCTCTCATCGAGCATCTGCAATTCAGATGATGTC	909
QY	913	AGAGATGCTGCTGCTTCAGACTGGCAGGGTCTGATGTTTCTGATGTTTCTGATGTTT	972
DB	910	AGAGATGCTGCTGCTTCAGACTGGCAGGGTCTGATGTTTCTGATGTTTCTGATGTTT	969
QY	973	AACCACTGCTGTTGGACTCGAGAGAAATTTCTAACTTTAAAGCTCCCGAGAGCATCTCAT	1032
DB	970	AACCACTGCTGTTGGACTCGAGAGAAATTTCTAACTTTAAAGCTCCCGAGAGCATCTCAT	1029
QY	1033	AAAGCTGTGCTCAATGGAAACATTTATGTTGGGTTGTTGGAGGATATATGTTCAACCTCA	1092
DB	1030	AAAGCTGTGCTCAATGGAAACATTTATGTTGGGTTGTTGGAGGATATATGTTCAACCTCA	1089



[illegible]

Db	1030	AAAGCTGTGCTCAATGAAACATTTATGTGGTGTGTTGGAGGATATATGTTCAACCACTCA	1089
Qy	1093	GATTATAACATGCTTCTAGCGTATGACCTTGTCTTAGCGAGTGGCTTCCACTAAACCGT	1152
Db	1090	GATTATAACATGCTTCTAGCGTATGACCTTGTCTTAGCGAGTGGCTTCCACTAAACCGT	1149
Qy	1153	TCTGTGAACAAATGCTGTGTTGATATGTTGTCATCTTTTGGCATATACAGAGATAAAAT	1212
Db	1150	TCTGTGAACAAATGCTGTGTTGATATGTTGTCATCTTTTGGCATATACAGAGATAAAAT	1209
Qy	1213	TACATGTATGAGGAGAAATTTGATCCAACTGGGAAATGTGACCAATGAGTTGAGAGTTT	1272
Db	1210	TACATGTATGAGGAGAAATTTGATCCAACTGGGAAATGTGACCAATGAGTTGAGAGTTT	1269
Qy	1273	CACATTCATATGAGTCATGGGTGTGTTGACCCCTAAGCCAAAGAGCAGTATGCAATG	1332
Db	1270	CACATTCATATGAGTCATGGGTGTGTTGACCCCTAAGCCAAAGAGCAGTATGCAATG	1329
Qy	1333	GTTGGGCACTCTGCACACATGTTACACGAAATGGCGAGTGTGCTGCTGTCATC	1392
Db	1330	GTTGGGCACTCTGCACACATGTTACACGAAATGGCGAGTGTGCTGCTGTCATC	1389
Qy	1393	TTTGGTCACTGCCCTCTCTATGATATATAGCAATGTGAGGAATATGATTTGGATAAG	1452
Db	1390	TTTGGTCACTGCCCTCTCTATGATATATAGCAATGTGAGGAATATGATTTGGATAAG	1449
Qy	1453	AACATGAGATATATACACACCGAGGTGCCCTTGTGCAAGGGGTTACGGCCATAGC	1512
Db	1450	AACATGAGATATATACACACCGAGGTGCCCTTGTGCAAGGGGTTACGGCCATAGC	1509
Qy	1513	AGTGTTTACCAACATAGGACCGAGGCTTATAGCTTTCATGGTGGCTACAGGCTTTCAGT	1572
Db	1510	AGTGTTTACCAACATAGGACCGAGGCTTATAGCTTTCATGGTGGCTACAGGCTTTCAGT	1569
Qy	1573	GCCAAATAGTACCGGCTTGCAGATGATCTCTACGATATGATGTGGATATCCAGATGTGG	1632
Db	1570	GCCAAATAGTACCGGCTTGCAGATGATCTCTACGATATGATGTGGATATCCAGATGTGG	1629
Qy	1633	ACCAATTTAAGGACAGCCGATTTTCCGTTACTTGCACACACAGCTGTGATGAGTGA	1692
Db	1630	ACCAATTTAAGGACAGCCGATTTTCCGTTACTTGCACACACAGCTGTGATGAGTGA	1689
Qy	1693	ACCAATGCTGTGTTTGGGGGAAACACACAAATGACACATCTATGAGCCATGCGCCAAA	1752
Db	1690	ACCAATGCTGTGTTTGGGGGAAACACACAAATGACACATCTATGAGCCATGCGCCAAA	1749
Qy	1753	TGCTTCTCTCAGATTTCAATGGCCCTATGACATTTGCCCTGTGACCGCTGCTGCTTCCC	1812
Db	1750	TGCTTCTCTCAGATTTCAATGGCCCTATGACATTTGCCCTGTGACCGCTGCTGCTTCCC	1809
Qy	1813	AGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTTACACACAGC	1872
Db	1810	AGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTTACACACAGC	1869
Qy	1873	ACCATGTATGTTTCCGTTGTTTCAATAGTCTCTCTCAGCGACATCTCTGGTATTCAAC	1932
Db	1870	ACCATGTATGTTTCCGTTGTTTCAATAGTCTCTCTCAGCGACATCTCTGGTATTCAAC	1929
Qy	1933	TCGGAACAGTGTGATGCGATCGAGTGAAGCGCTTGTGTTTACGACGAGCCTGTTATT	1992
Db	1930	TCGGAACAGTGTGATGCGATCGAGTGAAGCGCTTGTGTTTACGACGAGCCTGTTATT	1989
Qy	1993	CGGTGTGTGTGGAACACAGAGGTCTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAA	2052
Db	1990	CGGTGTGTGTGGAACACAGAGGTCTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAA	2049
Qy	2053	CAAGAGAAAGTTTAAATCAGATGTTTTCGAAAGAACTCTTGACCAATGACAGATGT	2112
Db	2050	CAAGAGAAAGTTTAAATCAGATGTTTTCGAAAGAACTCTTGACCAATGACAGATGT	2109
Qy	2113	GACCAGCACAGATTTGTACAGCTGTACAGCCAAACCAATGACTGCCACTGCTGTCAT	2172

Db	2110	GACCAGCACAGATTTGTTCAGCTGCACAGCCCAACCAATGACTGCCACTGCTGCAAT	2169
Qy	2173	GACCAATTTGTCCCGAGAAACACAGCTGTCTCAGAGCCAGATCTCCATTTTGTAGTAT	2232
Db	2170	GACCAATTTGTCCCGAGAAACACAGCTGTCTCAGAGCCAGATCTCCATTTTGTAGTAT	2229
Qy	2233	GAGAAATTCGCCCAAGGATAACCCCATGTTACTACTGTAAACAAGAGACAGCTGCAGGAGC	2292
Db	2230	GAGAAATTCGCCCAAGGATAACCCCATGTTACTACTGTAAACAAGAGACAGCTGCAGGAGC	2289
Qy	2293	TGTGCCCTGGAGACCAAGACTGCCAGTGGAGGCCCGGAATCAGGAGTGCATTTGCCCTGCC	2352
Db	2290	TGTGCCCTGGAGACCAAGACTGCCAGTGGAGGCCCGGAATCAGGAGTGCATTTGCCCTGCC	2349
Qy	2353	GAAATATCTGTGGCATTTGGCTGGCAATTTGGTTGGAACTCATGTTTGAATTAATCTACT	2412
Db	2350	GAAATATCTGTGGCATTTGGCTGGCAATTTGGTTGGAACTCATGTTTGAATTAATCTACT	2409
Qy	2413	GCCAAAGGAGAAATTAAGCAATGCTAAATTTGTTCTGTAGGAAACCAATGCCCTTTTGGCT	2472
Db	2410	GCCAAAGGAGAAATTAAGCAATGCTAAATTTGTTCTGTAGGAAACCAATGCCCTTTTGGCT	2469
Qy	2473	TCTCTTAAACCCAGAGAGGTAGAAATTTGCTCTTAAAGAGCTGCCGAATTAATGCAGTCA	2532
Db	2470	TCTCTTAAACCCAGAGAGGTAGAAATTTGCTCTTAAAGAGCTGCCGAATTAATGCAGTCA	2529
Qy	2533	TCTCAGAGCATGCTCCAAAGCTCACTTAACCCCATGGGTCCGGCTTCGGAAGATCAATGTG	2592
Db	2530	TCTCAGAGCATGCTCCAAAGCTCACTTAACCCCATGGGTCCGGCTTCGGAAGATCAATGTG	2589
Qy	2593	TCCTACTGGTGTGGGAAGATATGTCCTTCAATGTTTCAAAATAGTTTACTACAGTGGATGCG	2652
Db	2590	TCCTACTGGTGTGGGAAGATATGTCCTTCAATGTTTCAAAATAGTTTACTACAGTGGATGCG	2649
Qy	2653	TCTGAGCCAGTGTGCTGGAATTTCTGCTGGAATTTTATCAGAACCCAGTACTCGGGAGCTG	2712
Db	2650	TCTGAGCCAGTGTGCTGGAATTTCTGCTGGAATTTTATCAGAACCCAGTACTCGGGAGCTG	2709
Qy	2713	AAGGCTGCAACCTGCATCAACCCACTCAATGCTGTGTGAAAGGCTTCGAAACCAAC	2772
Db	2710	AAGGCTGCAACCTGCATCAACCCACTCAATGCTGTGTGAAAGGCTTCGAAACCAAC	2769
Qy	2773	AGTGTAGAGCTGCGGAGACACCATGCTTGGAGACAGCATGTGGAGATGTCACCAAGC	2832
Db	2770	AGTGTAGAGCTGCGGAGACACCATGCTTGGAGACAGCATGTGGAGATGTCACCAAGC	2829
Qy	2833	GGCAGCTCTGAGTGTGCTGTGAGCAACATGAGCAGTGTGGAATCCCAATGCCCTAT	2892
Db	2830	GGCAGCTCTGAGTGTGCTGTGAGCAACATGAGCAGTGTGGAATCCCAATGCCCTAT	2889
Qy	2893	GTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACTTGCCTTCCCT	2952
Db	2890	GTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACTTGCCTTCCCT	2949
Qy	2953	GAAATTTGTTCAGGCTACTGTACTCTGTAGTCTTGGAGCAACAGGCTGTGCTGAGTCC	3012
Db	2950	GAAATTTGTTCAGGCTACTGTACTCTGTAGTCTTGGAGCAACAGGCTGTGCTGAGTCC	3009
Qy	3013	TGTACTGATCCAGCAATACTGCGCAAGGAAATGATAGAGGGTTCCTTATAAGAGACCA	3072
Db	3010	TGTACTGATCCAGCAATACTGCGCAAGGAAATGATAGAGGGTTCCTTATAAGAGACCA	3069
Qy	3073	GTGAAGATGCTTCCGAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATTCC	3132
Db	3070	GTGAAGATGCTTCCGAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATTCC	3129
Qy	3133	AGCATGTGTCTAGAGGACAGCATCACTGCTTCTTCAATTCAGTGTCCAGCTTGGCCAA	3192
Db	3130	AGCATGTGTCTAGAGGACAGCATCACTGCTTCTTCAATTCAGTGTCCAGCTTGGCCAA	3189
Qy	3193	TGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACC	3252
Db	3190	TGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACC	3249



QY 3253 ACAGCAAGCACTGGAGACCTGCATATCTGGCTTCTACGGTGATCCACCAATGGAGG 3312  
Db 3250 ACAGCAAGCACTGGAGACCTGCATATCTGGCTTCTACGGTGATCCACCAATGGAGG 3309  
QY 3313 AAATGTCCAGCATCAAGTGCATGGGACGCGTCTCTGTGCAACCAACACGGGCAAG 3372  
Db 3310 AAATGTCCAGCATCAAGTGCATGGGACGCGTCTCTGTGCAACCAACACGGGCAAG 3369  
QY 3373 TGCCTTCTCCACCAAGGGGCTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAT 3432  
Db 3370 TGCCTTCTCCACCAAGGGGCTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAT 3429  
QY 3433 CGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATATCTCTTTATTGACTATCAG 3492  
Db 3430 CGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATATCTCTTTATTGACTATCAG 3489  
QY 3493 TTACCTTTAGTCTATCCAGGAAGATGATGCTATTTACAGTATCAATTTTGGCT 3552  
Db 3490 TTACCTTTAGTCTATCCAGGAAGATGATGCTATTTACAGTATCAATTTTGGCT 3549  
QY 3553 ACTCTGACCAACCAACAGGATTTGCAATGTTTCATCAATGCTCCAGAAATTTCAAC 3612  
Db 3550 ACTCTGACCAACCAACAGGATTTGCAATGTTTCATCAATGCTCCAGAAATTTCAAC 3609  
QY 3613 CTCACATACCTGGCTGCGAGTTTCTACGTGGAACCCAGGCTGGAGAGATGCGCT 3672  
Db 3610 CTCACATACCTGGCTGCGAGTTTCTACGTGGAACCCAGGCTGGAGAGATGCGCT 3669  
QY 3673 GTTCTTTCAAAACCAACATTAAGGATCAAGAGATGTTCTCTAATGAGAGTTTGAT 3732  
Db 3670 GTTCTTTCAAAACCAACATTAAGGATCAAGAGATGTTCTCTAATGAGAGTTTGAT 3729  
QY 3733 TTTGCGAACCAACCAATATACATTTCTTTGTTTATGTCAGTAAATTTCACTGGCCCATC 3792  
Db 3730 TTTGCGAACCAACCAATATACATTTCTTTGTTTATGTCAGTAAATTTCACTGGCCCATC 3789  
QY 3793 AAAATTGAG 3801  
Db 3790 AAAATTGAG 3798

RESULT 4  
AF034957  
LOCUS AF034957 3597 bp mRNA linear PRI 30-SEP-1998  
DEFINITION Homo sapiens secreted T cell activation protein Attractin  
(attractin) mRNA, complete cds.  
ACCESSION AF034957  
VERSION AF034957.1 GI:3676346  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3597)  
Duke-Cohan,J.S., Gu,J., McLaughlin,D.F., Xu,Y., Freeman,G.J. and  
Schlossman,S.F.  
TITLE Attractin (DPPT-L), a member of the CUB family of cell adhesion and  
guidance proteins, is secreted by activated human T lymphocytes and  
modulates immune cell interactions  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)  
MEDLINE 98409658  
PUBMED 9736737  
REFERENCE 2 (bases 1 to 3597)  
Duke-Cohan,J.S., Gu,J., Ao,Z., McLaughlin,D.F., Freeman,G.J. and  
Schlossman,S.F.  
AUTHORS Direct Submission  
TITLE Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS,  
JOURNAL Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,  
USA  
FEATURES  
Location/Qualifiers  
1..3597  
/organism="Homo sapiens"

/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..3597  
/gene="attractin"  
1..3597  
/gene="attractin"  
/note="secreted T cell activation protein; Method:  
conceptual translation with partial peptide sequencing;  
dppt-l"  
/codon\_start=1  
/product="Attractin"  
/protein\_id="AAC61902.1"  
/db\_xref="GI:3676347"  
/translation="MVAAAATAEARLRRTAATAALAGRSGPHGVNRCNPGTGC  
VCPAGVGEQCHQGRFRLTSSGFTVTDGNGYKTKTCTWLI EGGPNIRLRFNH  
FATEGSDHLYVVDGDSIYALVAAGSLI VPERDNETVEPVATSGADIPHCTDNG  
AYNLGNITYSFDMCPNCSGEGECKISNSDTVECESENWKEACDIPHCTDNG  
FPHRIGNSDVRGCSGSDQWGFCSVPVPANOSFWTREYSNLKLPRAHKAAYNG  
NIMWGYGFNHSNDYNVLAIDLASREWLPNRSVNVVRYVGHSLALYKDKLIYTG  
GKIDSTGNVTNELRVFHNESWLLTPKAQEYAVVGHSAHIVLKNRVMVLIYEG  
HCLYGYISNVQYDLDKNTWSILHTQALVQGGYGHSSVYDHRTRALYVHGGYKAS  
ANKRYLADLVRYDVTOMTILKDSRFFRYLHTAVIVSGTLMVFGNTNHDGSMHS  
AKFSSDDPMAYDIACDRWSVLPDPSITMMSDOLAI PAVLHNSWVYVFGNSLLSDI  
LVFTSECCDAHSEAACLAAGPGIRCVWNTGSSQCI SKALATDEBEKLSCEPKET  
LDHRCDOHTICYSTANDHCNDHCNPRNSHSCSEGGQISIFRYENCPKDNPMYC  
NKTSCRSKALDQNCQEPNQCIALPENICGILGHLVGNLSCLITAKENYDNAL  
FCRNHALLASLTITKQVRFVQLIRIQSSQMSKLTLPWGLRKNINYSYCWEDM  
SPFNLSLQWMPSEPSDAGFCGILSEPSRGLKAATCINPLNGSVCRPAHSAKQR  
TPCALRTACGDCSTSGSSECMWCSNMKOCVDSNAVASFPFQCMVMTMGTCPDENS  
GYCTSCHELEGCGNCTDPSNTCKGKICISYKGVKPMPSOAPTGNFYQPLNLSM  
CLESRYNWSFIHCPACQNGHCKINQSCICEKENLITGKHCTCISGYPDFTNGG  
CKPCKNGHSLCNTNKGFCITTKGKGDEQCLCEVENRYQNPURGICITYLLID  
YQFTFVSQDDRYITAINFVATPEQNRDLDMFINASKNFNLNITWAASFSQTAG  
BEMPVSKNTNIKEYKDSFSNEKEDFRNHPNITFEFVYVSNFTWPIKIQVQTEQ"

ORIGIN  
Query Match 71.0%; Score 3045; DB 9; Length 3597;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3485; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
QY 312 CTGTGTCAACGGCGGTGCTGCAACCCCTGGCACCGGCGAGTGCCTGCGCCCGCGGCTG 371  
Db 90 CTGTGTCAACGGCGGTGCTGCAACCCCTGGCACCGGCGAGTGCCTGCGCCCGCGGCTG 149  
QY 372 GGTGGGCGAGCAATGCCAGCACTCGGGGGCGCTTCAGACTAACTGATCTTCGGGTT 431  
Db 150 GGTGGGCGAGCAATGCCAGCACTCGGGGGCGCTTCAGACTAACTGATCTTCGGGTT 209  
QY 432 TGTGACAGATGACCTGGAATTAATAACAAACGAAAGTGACGCTGGCTCATTTGAAGG 491  
Db 210 TGTGACAGATGACCTGGAATTAATAACAAACGAAAGTGACGCTGGCTCATTTGAAGG 269  
QY 492 ACAGCAAAATAGATAATGAGACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTGGGA 551  
Db 270 ACAGCAAAATAGATAATGAGACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTGGGA 329  
QY 552 CCATTATATCTTTATGATGGGACTCAATTTATGACCGCTAGTCTGCTGCTATTAGTGG 611  
Db 330 CCATTATATCTTTATGATGGGACTCAATTTATGACCGCTAGTCTGCTGCTATTAGTGG 389  
QY 612 CCTCATTTCTCTCAGAGAGATGCAATGAGCTGTCCCTGAGGTTGTGCCACATCAGG 671  
Db 390 CCTCATTTCTCTCAGAGAGATGCAATGAGCTGTCCCTGAGGTTGTGCCACATCAGG 449  
QY 672 TTATGCTTCTGCTGATTTTATGATGCTGCTTAAATTTAGTCTGATTTATATTATAC 731  
Db 450 TTATGCTTCTGCTGATTTTATGATGCTGCTTAAATTTAGTCTGATTTATATTATAC 509  
QY 732 TTACAGTTTTGATATGTCTCAAAATACTGTCTCAGGCGGAGGAGGTGAAGTACAGTAA 791  
Db 510 TTACAGTTTTGATATGTCTCAAAATACTGTCTCAGGCGGAGGAGGTGAAGTACAGTAA 569  
QY 792 TAGCAGGAAACTGTTGAATGTGAATGTTCTGAAAACCTGGAAGGTGAAGCATGTGACAT 851

570 TAGCAGGATCTGTTGAAATGTAATGTTCTGAAACTGAAAGGTGAAGCATGTGACAT 629 Db  
852 TCCTCACTGTACAGACAACCTGTGGTCTTCTCATCGAGGATCTGCAATTCAGATGATGT 911 Qy  
630 TCCTCACTGTACAGACAACCTGTGGTCTTCTCATCGAGGATCTGCAATTCAGATGATGT 689 Db  
912 CAGAGGATGTCCTGCTTCTCAGACTGGCAGGGTCTCTGGATGTTCAAGTTCCTGTACCAGC 971 Qy  
690 CAGAGGATGTCCTGCTTCTCAGACTGGCAGGGTCTCTGGATGTTCAAGTTCCTGTACCAGC 749 Db  
972 TAACCACTCAATTTGGACTCGAGAGGAATATTTAACTTAAAGCTCCCGAGACATCTCA 1031 Qy  
750 TAACCACTCAATTTGGACTCGAGAGGAATATTTAACTTAAAGCTCCCGAGACATCTCA 809 Db  
1032 TAAAGCTGTGGTCAATCGGAACATTAATGCTGGTGTGGAGGATATATGTTCAACCACTC 1091 Qy  
810 TAAAGCTGTGGTCAATCGGAACATTAATGCTGGTGTGGAGGATATATGTTCAACCACTC 869 Db  
1092 AGATTATAAATGCTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCG 1151 Qy  
870 AGATTATAAATGCTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCG 929 Db  
1152 TTCTGTGAACAATGTGGTGTAGATATGTCATCTTTGGCAATATACAGGATAAAT 1211 Qy  
930 TTCTGTGAACAATGTGGTGTAGATATGTCATCTTTGGCAATATACAGGATAAAT 989 Db  
1212 TTACATGCTATGGAGGAAAAATTTGATCCAACTGGGAATGTCAACATGAGTGTAGAGTTTT 1271 Qy  
990 TTACATGCTATGGAGGAAAAATTTGATCCAACTGGGAATGTCAACATGAGTGTAGAGTTTT 1049 Db  
1272 TCACATTCATAAATGAGTCATGGGTGTTGTTGACCCCTTAAGCCAAAGGAGCAGTATGAGT 1331 Qy  
1050 TCACATTCATAAATGAGTCATGGGTGTTGTTGACCCCTTAAGCCAAAGGAGCAGTATGAGT 1109 Db  
1332 GGTGGGCACTCTGCACACATGTTTACACTGAGAAATGGCCGAGTGTCTATGCTGTCTCAT 1391 Qy  
1110 GGTGGGCACTCTGCACACATGTTTACACTGAGAAATGGCCGAGTGTCTATGCTGTCTCAT 1169 Db  
1392 CTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTGGATAA 1451 Qy  
1170 CTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTGGATAA 1229 Db  
1452 GAACACATGAGTATATACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAG 1511 Qy  
1230 GAACACATGAGTATATACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATAG 1289 Db  
1512 CAGTGTGTTACGACCATAGGACCCAGGCCCTATACGTTTCATGGTGGCTACAAGGCTTTTCAG 1571 Qy  
1290 CAGTGTGTTACGACCATAGGACCCAGGCCCTATACGTTTCATGGTGGCTACAAGGCTTTTCAG 1349 Db  
1572 TGCCAATAAGTACCCGGCTTGAGATGATCTCTACCGATATGATGTGGATACCCAGATGTG 1631 Qy  
1350 TGCCAATAAGTACCCGGCTTGAGATGATCTCTACCGATATGATGTGGATACCCAGATGTG 1409 Db  
1632 GACCATTCCTTAAGACACCGAATTTTCCGTTATCTGCACACAGCTGTGATGTAGTGG 1591 Qy  
1410 GACCATTCCTTAAGACACCGAATTTTCCGTTATCTGACACAGCTGTGATGTAGTGG 1469 Db  
1692 AACCATGCTGGTGTGTTGGGGAACACACACAATGACACATCTATGAGCCATGCGCCAA 1751 Qy  
1470 AACCATGCTGGTGTGTTGGGGAACACACACAATGACACATCTATGAGCCATGCGCCAA 1529 Db  
1752 ATGCTTCTCTTCAATTTCAATGGCGTATGACATTTGCCCTGTGACCGCTGGTCACTTCC 1811 Qy  
1530 ATGCTTCTCTTCAATTTCAATGGCGTATGACATTTGCCCTGTGACCGCTGGTCACTTCC 1589 Db  
1812 CAGACCTCATCTCCACCATGATGTCAACAGATTTGGCCATT - CAGCAGTCTTACACAACA 1870 Qy  
1590 CAGACCTCAT - TCACCATGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAACA 1648 Db  
1871 GCACCATGATGTGTGGTGGTTCATATAGTCTCTCTCCAGCGCATCTCTGGTATTCAT 1930 Qy

1649 GCACCATGATGTGTTCGGTGGTTCCTCAATAGTCTCCTCCTCAGCGACATCTCTGGTATTCAT 1708 Db  
1931 CCTCCGACACATGTGATCGGCATCGAGTGAAGCCGCTTGTGTAGCAGCAGACCTGTGTA 1990 Qy  
1709 CCTCCGACACATGTGATCGGCATCGAGTGAAGCCGCTTGTGTAGCAGCAGACCTGTGTA 1768 Db  
1991 TTCGTGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATG 2050 Qy  
1769 TTCGTGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATG 1828 Db  
2051 AACAGAGAAAAGTTAAATCAGAAATGTTTTCCAAAAGAACTCTTGACCATGACAGAT 2110 Qy  
1829 AACAGAGAAAAGTTAAATCAGAAATGTTTTCCAAAAGAACTCTTGACCATGACAGAT 1888 Db  
2111 GTGACCAACACACAGATTTTACAGCTGTACAGCCAAACACCAATGACTGCCACTGTGCA 2170 Qy  
1889 GTGACCAACACACAGATTTTACAGCTGTACAGCCAAACACCAATGACTGCCACTGTGCA 1948 Db  
2171 ATGACCAATTTGTCCTCCAGGAACACACAGTGTCTGAGAGGCCAGATCTCCATTTTAGGT 2230 Qy  
1949 ATGACCAATTTGTCCTCCAGGAACACACAGTGTCTGAGAGGCCAGATCTCCATTTTAGGT 2008 Db  
2231 ATGAGAAATTTGCCCAAGGATTAACCCATGTACTGTAAACAAGAACCAAGCTGCAGGA 2290 Qy  
2009 ATGAGAAATTTGCCCAAGGATTAACCCATGTACTGTAAACAAGAACCAAGCTGCAGGA 2068 Db  
2291 GCTGTGCCCTGGACAGAACTGCCAGTGGAGCCGCCGAATCAGGAGTGCATTTGCCCTGC 2350 Qy  
2069 GCTGTGCCCTGGACAGAACTGCCAGTGGAGCCGCCGAATCAGGAGTGCATTTGCCCTGC 2128 Db  
2351 CCGAAAATATCTGTGGCATTTGGTGGAACTCATGTTTGAATAATCTA 2410 Qy  
2129 CCGAAAATATCTGTGGCATTTGGTGGAACTCATGTTTGAATAATCTA 2188 Db  
2411 CTGCCAAGAGAAATTTATGACAAATGCTAAATTTCTGTAGGAACCAACAATGCCCTTTGG 2470 Qy  
2189 CTGCCAAGAGAAATTTATGACAAATGCTAAATTTCTGTAGGAACCAACAATGCCCTTTGG 2248 Db  
2471 CTTCTCTTAACACCAGAAAGAGTAGAATTTGCTTAAAGCAGCTGCGAATTAATGCAATG 2530 Qy  
2249 CTTCTCTTAACACCAGAAAGAGTAGAATTTGCTTAAAGCAGCTGCGAATTAATGCAATG 2308 Db  
2531 CATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGTCCGCCCTCGGAAGATCAATG 2590 Qy  
2309 CATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGTCCGCCCTCGGAAGATCAATG 2368 Db  
2591 TGTCTCTACTGTGTGGGAAGATATGTCCCATTTACAAATAGTTTATCTACAGTGGATGC 2650 Qy  
2369 TGTCTCTACTGTGTGGGAAGATATGTCCCATTTACAAATAGTTTATCTACAGTGGATGC 2428 Db  
2651 CGTCTGAGCCAGTGTGATGCTGGATTTCTGGAAATTTTATCAGAACCCAGTACTCGGGAC 2710 Qy  
2429 CGTCTGAGCCAGTGTGATGCTGGATTTCTGGAAATTTTATCAGAACCCAGTACTCGGGAC 2488 Db  
2711 TGAAGGCTGCAACCTGCATCAACCCCATCAATGAGTGTCTGTGAAAGGCTCGCAACCC 2770 Qy  
2489 TGAAGGCTGCAACCTGCATCAACCCCATCAATGAGTGTCTGTGAAAGGCTCGCAACCC 2548 Db  
2771 ACAGTCTTAAGCAGTGTCCGGAACACCATGTGCTTGAAGACAGCATGTGGAGATGCAACA 2830 Qy  
2549 ACAGTCTTAAGCAGTGTCCGGAACACCATGTGCTTGAAGACAGCATGTGGAGATGCAACA 2608 Db  
2831 GGGGAGCTCTCAGTGTGATGCTGGTGCAGCAACATGAAGCAGTGTGTGAACTCCAATGCC 2890 Qy  
2609 GGGGAGCTCTCAGTGTGATGCTGGTGCAGCAACATGAAGCAGTGTGTGAACTCCAATGCC 2668 Db  
2891 ATGTGCCCTCTTCCCTTTTGGCCAGTGTGAAATGGAATGGAATGAGTACATGAGACACTGCCCC 2950 Qy  
2669 ATGTGCCCTCTTCCCTTTTGGCCAGTGTGAAATGGAATGGAATGAGTACATGAGACACTGCCCC 2728 Db  
2951 CTGAAAATTTGTCAGGCTACTCTGCTGTAGTCAATTTGTTGAGCAACACAGGCTGTGGCT 3010 Qy  
2729 CTGAAAATTTGTCAGGCTACTCTGCTGTAGTCAATTTGTTGAGCAACACAGGCTGTGGCT 2788 Db

QY	3011	GGTGTACTGATCCCGAGCAATACTGCGAAAGGAAATGATAGAGGGTTCCTATTAAGGAC	3070
Db	2789	GGTGTACTGATCCCGAGCAATACTGCGAAAGGAAATGATAGAGGGTTCCTATTAAGGAC	2848
QY	3071	CAGTGAAGATGCTTCGCAAGCCCTCAGAGAAATTTCTATCCACGCCCTGCTCAATT	3130
Db	2849	CAGTGAAGATGCTTCGCAAGCCCTCAGAGAAATTTCTATCCACGCCCTGCTCAATT	2908
QY	3131	CCAGCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTTCAATTCAGTCTCCAGCTTGGC	3190
Db	2909	CCAGCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTTCAATTCAGTCTCCAGCTTGGC	2968
QY	3191	AATGCAAGGGCCACAGTAAATGCAATCAATCAGACATCTGTGAGAGTGTGAGAACCTGA	3250
Db	2969	AATGCAAGGGCCACAGTAAATGCAATCAATCAGACATCTGTGAGAGTGTGAGAACCTGA	3028
QY	3251	CCAGGCAAGCACTGCGAGACCTGCAATCTGGCTTCTACGGTGTATCCCAACCAATGGAG	3310
Db	3029	CCAGGCAAGCACTGCGAGACCTGCAATCTGGCTTCTACGGTGTATCCCAACCAATGGAG	3088
QY	3311	GGAAATGTCAGCCATGCAAGTGCATGAGGCGACCGCTCTCTGTGCAACACCAACACGGGCA	3370
Db	3089	GGAAATGTCAGCCATGCAAGTGCATGAGGCGACCGCTCTCTGTGCAACACCAACACGGGCA	3148
QY	3371	AGTGTCTCTGCACCAACCAAGGGCGTCAAGGGGACGAGTGCAGCTATGTAGGTAGAAA	3430
Db	3149	AGTGTCTCTGCACCAACCAAGGGCGTCAAGGGGACGAGTGCAGCTATGTAGGTAGAAA	3208
QY	3431	ATCGATACCAAGGAAACCTCTCAGAGGAAATGTTATTAATCTCTTTTATTTGACTATC	3490
Db	3209	ATCGATACCAAGGAAACCTCTCAGAGGAAATGTTATTAATCTCTTTTATTTGACTATC	3268
QY	3491	AGTTCACCTTTAGTCTATCCAGAGAGATGATCGCTATTATACAGAGCTATCAATTTTGTGG	3550
Db	3269	AGTTCACCTTTAGTCTATCCAGAGAGATGATCGCTATTATACAGAGCTATCAATTTTGTGG	3328
QY	3551	CTACTCTCTGAGCAACAAACAGGGATTTGACATGTTTCATCAATGCCCTCCAGAAATTTCA	3610
Db	3329	CTACTCTCTGAGCAACAAACAGGGATTTGACATGTTTCATCAATGCCCTCCAGAAATTTCA	3388
QY	3611	ACCTCAACATCACTGGGCTCCAGTTTCTCAGCTGGAAACCGAGCTGGAGAGAGATGC	3670
Db	3389	ACCTCAACATCACTGGGCTCCAGTTTCTCAGCTGGAAACCGAGCTGGAGAGAGATGC	3448
QY	3671	CTGTGTTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTG	3730
Db	3449	CTGTGTTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTG	3508
QY	3731	ATTTTCGCAACCCCAAAATATCACTTCTTTTGTATGTCAGTAATTTCACTGGCCCA	3790
Db	3509	ATTTTCGCAACCCCAAAATATCACTTCTTTTGTATGTCAGTAATTTCACTGGCCCA	3568
QY	3791	TCAAAATTCAG 3801	
Db	3569	TCAAAATTCAG 3579	

RESULT 5  
AR164818  
LOCUS AR164818 2625 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 18 from patent US 6274339.  
ACCESSION AR164818  
VERSION AR164818.1 GI:16238029  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 2625)  
Moore,K. and Nagle,D.Lynn.  
AUTHORS Methods and compositions for the diagnosis and treatment of body  
TITLE weight disorders, including obesity  
JOURNAL Patent: US 6274339-A 18 14-AUG-2001.

FEATURES		Location/Qualifiers	
source		1..2625	
		/organism="unknown"	
		/mol_type="unassigned DNA"	
ORIGIN			
Query Match		42.8%; Score 1837; DB 6; Length 2625;	
Best Local Similarity		99.8%; Pred. No. 0;	
Matches 2037; Conservative		0; Mismatches 4; Indels 0; Gaps 0;	
Qy	313	TGTGTCAAACGGCGGTGCTGCAACCCCTGGCACCGGCCAGTGGCTGCTGCCCGCGGTGG	372
Db	310	TGTGTCAAACGGCGGTGCTGCAACCCCTGGCACCGGCCAGTGGCTGCTGCCCGCGGTGG	369
Qy	373	GTGGGCGAGCAATGCCAGCACTGGCGGGGCCGCTTCAGACTAACTGGATCTCTCGGTTT	432
Db	370	GTGGGCGAGCAATGCCAGCACTGGCGGGGCCGCTTCAGACTAACTGGATCTCTCGGTTT	429
Qy	433	GTGACAGATGGACCTGGAAATTTATAAATACAAACGAAAGTGCACGTGGCTCAATTGAAGA	492
Db	430	GTGACAGATGGACCTGGAAATTTATAAATACAAACGAAAGTGCACGTGGCTCAATTGAAGA	489
Qy	493	CAGCCAAATAGAAATAATAGACTTCGTTTCAATCAATTTTGTCTACAGAGTGTAGTTGGGAC	552
Db	490	CAGCCAAATAGAAATAATAGACTTCGTTTCAATCAATTTTGTCTACAGAGTGTAGTTGGGAC	549
Qy	553	CATTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGC	612
Db	550	CATTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGC	609
Qy	613	CTCATTGTTTCTGAGAGAGATGGCAATGAGACTGTCCTCGAGGTTGTTGCCACATCAGGT	672
Db	610	CTCATTGTTTCTGAGAGAGATGGCAATGAGACTGTCCTCGAGGTTGTTGCCACATCAGGT	669
Qy	673	TATGCTTGTGTCATTTTTTTAGTATGCTGCTTATAATTTGACTGGATTTAATTAATTACT	732
Db	670	TATGCTTGTGTCATTTTTTTAGTATGCTGCTTATAATTTGACTGGATTTAATTAATTACT	729
Qy	733	TACAGTTTGTATGTCCTCAAAATAAATGCTCAGGCCGAGGAGAGTGAAGATCAGTAAT	792
Db	730	TACAGTTTGTATGTCCTCAAAATAAATGCTCAGGCCGAGGAGAGTGAAGATCAGTAAT	789
Qy	793	AGCAGCGAAACTGTTGAATGTGAATGTTCTGAAAATCTGAAAAGGTGAAGCATGTGACATT	852
Db	790	AGCAGCGATACTGTTGAATGTGAATGTTCTGAAAATCTGAAAAGGTGAAGCATGTGACATT	849
Qy	853	CCTCACTGTACAGACAATGTTGTTTCTCATCGAGGCATCTGCAATTTCAAGTGAATGC	912
Db	850	CCTCACTGTACAGACAATGTTGTTTCTCATCGAGGCATCTGCAATTTCAAGTGAATGC	909
Qy	913	AGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCTCGGATGTTTCAGTTCTCTGTACACGT	972
Db	910	AGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCTCGGATGTTTCAGTTCTCTGTACACGT	969
Qy	973	AACCACTCATTTTGGACTCCGAGAGAAATTTCTTAACCTTAAAGCTCCCGAGAGCATCTCAT	1032
Db	970	AACCACTCATTTTGGACTCCGAGAGAAATTTCTTAACCTTAAAGCTCCCGAGAGCATCTCAT	1029
Qy	1033	AAAGCTGTGGTCAATGGAAAATTTATGTTGGGTTGTTGGAGGATATATGTTCAACCACTCA	1092
Db	1030	AAAGCTGTGGTCAATGGAAAATTTATGTTGGGTTGTTGGAGGATATATGTTCAACCACTCA	1089
Qy	1093	GATTATPAAATGGTTCACGGTATGACCTTGGCTTCTAGGGAGTGGCTTCCACTAAACCGT	1152
Db	1090	GATTATPAAATGGTTCACGGTATGACCTTGGCTTCTAGGGAGTGGCTTCCACTAAACCGT	1149
Qy	1153	TCTGTGAAACAATGTTGTTTGTAGATATGCTCAATTTCTTTGGCAATTATACAAGGATAAAAT	1212
Db	1150	TCTGTGAAACAATGTTGTTTGTAGATATGCTCAATTTCTTTGGCAATTATACAAGGATAAAAT	1209
Qy	1213	TACATGTATGAGGAAAATTTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTT	1272
Db	1210	TACATGTATGAGGAAAATTTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTT	1269

```
QY 1273 CACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGCGAAGAGCAGATATGCAAGT 1332
Db 1270 CACATTCATAATGAGTCATGGGTGTTGTTGACCCCTAAGCGAAGAGCAGATATGCAAGT 1329
QY 1333 GTTGGGCACCTCTGCACACATTTGTACATGAAGAATGGCGAGTGTCTATGCTGGTTCATC 1392
Db 1330 GTTGGGCACCTCTGCACACATTTGTACATGAAGAATGGCGAGTGTCTATGCTGGTTCATC 1389
QY 1393 TTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAAG 1452
Db 1390 TTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAAG 1449
QY 1453 AACACATGAGATATTTACACACCCAGGGTGCCTTTGTGCAAGGGGGTTACGGCCCATAGC 1512
Db 1450 AACACATGAGATATTTACACACCCAGGGTGCCTTTGTGCAAGGGGGTTACGGCCCATAGC 1509
QY 1513 AGTGTTTACACCATAGGACACAGGCCCTTATAGTTTCATGCTGGTGTCTACAGGCTTTTCAGT 1572
Db 1510 AGTGTTTACACCATAGGACACAGGCCCTTATAGTTTCATGCTGGTGTCTACAGGCTTTTCAGT 1569
QY 1573 GCCAATAAGTACCGGCTTGACAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGG 1632
Db 1570 GCCAATAAGTACCGGCTTGACAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGG 1629
QY 1633 ACCATTCTTAAGGACACCCGATTTTCCGTTTACTTGCACACAGCTGTGATGATGTGA 1692
Db 1630 ACCATTCTTAAGGACACCCGATTTTCCGTTTACTTGCACACAGCTGTGATGATGTGA 1689
QY 1693 ACCATGCTGTGTTTGGGGGAACACACACAATGACACATCTATGAGCCATGCGGCCAA 1752
Db 1690 ACCATGCTGTGTTTGGGGGAACACACACAATGACACATCTATGAGCCATGCGGCCAA 1749
QY 1753 TGTCTTCCTTCAGATTTTCATGGCCCTATGACATTCCTGTGACCGCTGGTCAGTCTTCCC 1812
Db 1750 TGTCTTCCTTCAGATTTTCATGGCCCTATGACATTCCTGTGACCGCTGGTCAGTCTTCCC 1809
QY 1813 AGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGCTTTACACACAGC 1872
Db 1810 AGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGCTTTACACACAGC 1869
QY 1873 ACCATGATGTGTCGGTGGTTTCAATAGTCTCTCTCCTCAGCAGCATCTCGTATTCACC 1932
Db 1870 ACCATGATGTGTCGGTGGTTTCAATAGTCTCTCTCCTCAGCAGCATCTCGTATTCACC 1929
QY 1933 TCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTTTATGACAGCAGCACTGGTATT 1992
Db 1930 TCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTTTATGACAGCAGCACTGGTATT 1989
QY 1993 CGGTGTGTGGAACACAGGGTCTCTCAGTGTATCTCGTGGGCGTGGCACTGATGAA 2052
Db 1990 CGGTGTGTGGAACACAGGGTCTCTCAGTGTATCTCGTGGGCGTGGCACTGATGAA 2049
QY 2053 CAAGAAGAAAAGTTAAATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGT 2112
Db 2050 CAAGAAGAAAAGTTAAATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGT 2109
QY 2113 GACACAGACACAGATTTGTACAGCTGTACAGCCCAACCAATGACTGCCACTGGTCAAT 2172
Db 2110 GACACAGACACAGATTTGTACAGCTGTACAGCCCAACCAATGACTGCCACTGGTCAAT 2169
QY 2173 GACCATTTGTGTCGCCAGGAACCAAGCTGTCTCAGAGGCCAGATCTCCATTTTAGGTAT 2232
Db 2170 GACCATTTGTGTCGCCAGGAACCAAGCTGTCTCAGAGGCCAGATCTCCATTTTAGGTAT 2229
QY 2233 GAGAAATGCCCAAGGATAACCCCATGTACTGTATCTGTAAACGAAGACAGCTGACGAGC 2292
Db 2230 GAGAAATGCCCAAGGATAACCCCATGTACTGTATCTGTAAACGAAGACAGCTGACGAGC 2289
QY 2293 TGTGCCCTGGACACAGAACTGCTGAGGAGCCCGGGAATCAGGAGTCAATTTGCCCTGCC 2352
Db 2290 TGTGCCCTGGACACAGAACTGCTGAGGAGCCCGGGAATCAGGAGTCAATTTGCCCTGCC 2349
```

```
QY 2353 G 2353
Db 2350 G 2350

RESULT 6
LOCUS AK000356
DEFINITION Homo sapiens cDNA FLJ20349 fis, clone HEP13957, highly similar to
ACCESSION AK000356
VERSION AK000356.1 GI:7020387
KEYWORDS cllgo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Ota,T.,
Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2. (bases 1 to 2578)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Department of Virology; Shirokane-dai, 4-6-1,
Minato-ku, Tokyo 108-8639, Japan (E-mail: fclcdna@ims.u-tokyo.ac.jp,
Tel:81-3-5449-5286, Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing; Research Association for Biotechnology; cDNA library
construction; 5'- & 3'-end one pass sequencing; Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency).
FEATURES
source
1. .2578
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP13957"
/cell_line="HepG2"
/cell_type="hepatoma"
/clone_lib="HEP"
/note="cloning vector pME18SFL3"
misc_feature
1. .2578
/note="highly similar to AB011120 Homo sapiens mRNA for
KIAA0548 protein"
ORIGIN
Query Match 39.6%; Score 1697; DB 9; Length 2578;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2594 CCTACTGTCGTCGGGAAGATATGCCCATTTACAAATAGTTTACTACAGTGCAGTCCGT 2653
Db 1 CCTACTGTCGTCGGGAAGATATGCCCATTTACAAATAGTTTACTACAGTGCAGTCCGT 60
QY 2654 CTGAGCCCACTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGA 2713
Db 61 CTGAGCCCACTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGA 120
QY 2714 AGGCTGCAACCTGTCATCAACCCACTCANTGGTAGTGTCTGTGAAGGCTGCAAAACCA 2773
Db 121 AGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACCA 180
QY 2774 GTGCTAAGCAGTGCAGGACACCATGTCCTTGAGACAGCATGTGAGATTGCACAGCG 2833
Db 181 GTGCTAAGCAGTGCAGGACACCATGTCCTTGAGACAGCATGTGAGATTGCACAGCG 240
```

```
QY 2834 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTTCCAATGCTTATG 2893
Db 241 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTTCCAATGCTTATG 300
QY 2894 TGGGCTCCCTTCCCTTTTGGCCAGTGTATGGATGGTATAGATGACATGACCTGCCCTG 2953
Db 301 TGGGCTCCCTTCCCTTTTGGCCAGTGTATGGATGGTATAGATGACATGACCTGCCCTG 360
QY 2954 AAAATTGTTCAAGGCTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGT 3013
Db 361 AAAATTGTTCAAGGCTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACTGT 420
QY 3014 GTACTGATCCAGCAATCTCTGGCAAGGAAATGCAATAGAGGTTCTTAAAGGACAG 3073
Db 421 GTACTGATCCAGCAATCTCTGGCAAGGAAATGCAATAGAGGTTCTTAAAGGACAG 480
QY 3074 TGAAGATGCTTCCAGAGCCCTACAGGAATTTCTATCCACAGCCCTGCTCAATTTCCA 3133
Db 481 TGAAGATGCTTCCAGAGCCCTACAGGAATTTCTATCCACAGCCCTGCTCAATTTCCA 540
QY 3134 GCATGTGTCTAGAGCAAGAGATACAACTGGTCTTTCAATCTGCTCCAGCTTGCCAAAT 3193
Db 541 GCATGTGTCTAGAGCAAGAGATACAACTGGTCTTTCAATCTGCTCCAGCTTGCCAAAT 600
QY 3194 GCAAGGCGCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAACTGACCA 3253
Db 601 GCAAGGCGCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAACTGACCA 660
QY 3254 CAGGCAAGCACTGGAGCACTGTCATATCTGGCTTTCTAGCGTGTATCCCAATGAGGGA 3313
Db 661 CAGGCAAGCACTGGAGCACTGTCATATCTGGCTTTCTAGCGTGTATCCCAATGAGGGA 720
QY 3314 AATGTACGCAATGAGTGAATGGGACCGGTCTCTGTGCAACACCAACAGCGGCAAGT 3373
Db 721 AATGTACGCAATGAGTGAATGGGACCGGTCTCTGTGCAACACCAACAGCGGCAAGT 780
QY 3374 GTTCTTGACCAACCAAGGGCGTCAAGGGGACAGTGCAGCTATGTAGGTAGAAAATC 3433
Db 781 GTTCTTGACCAACCAAGGGCGTCAAGGGGACAGTGCAGCTATGTAGGTAGAAAATC 840
QY 3434 GATACCAAGGAAACCTTCTCAGAGCAATGTTATATATCTCTTTTATTTGACTATCAGT 3493
Db 841 GATACCAAGGAAACCTTCTCAGAGCAATGTTATATATCTCTTTTATTTGACTATCAGT 900
QY 3494 TCACCTTTAGTCTATCCAGGAAGATGATCGTATTACAGCTATCAATTTTGTGGCTA 3553
Db 901 TCACCTTTAGTCTATCCAGGAAGATGATCGTATTACAGCTATCAATTTTGTGGCTA 960
QY 3554 CTCCTGACGAAACAAAACAGGAAATTTGACATGTTTCATCAATGCTCCCAAGATTTCAACC 3613
Db 961 CTCCTGACGAAACAAAACAGGAAATTTGACATGTTTCATCAATGCTCCCAAGATTTCAACC 1020
QY 3614 TCACATCAGCTGGCTGCGGCTTCTCAGTGGAAACCCAGGCTGGAGAGAGATGCGCTG 3673
Db 1021 TCACATCAGCTGGCTGCGGCTTCTCAGTGGAAACCCAGGCTGGAGAGAGATGCGCTG 1080
QY 3674 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATGTTTCTCTAATGAGAGTTTGATT 3733
Db 1081 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATGTTTCTCTAATGAGAGTTTGATT 1140
QY 3734 TTGCAACACCCAAATATCACTTTCTTTGTTATATGTCAGTAATTTTCACTGCCCATCA 3793
Db 1141 TTGCAACACCCAAATATCACTTTCTTTGTTATGTCAGTAATTTTCACTGCCCATCA 1200
QY 3794 AAATTTCAGATTGCTTCTCTCAGCACAGCAATTTTATGACCTGGTACAGTTCTTCTGTA 3853
Db 1201 AAATTTCAGATTGCTTCTCTCAGCACAGCAATTTTATGACCTGGTACAGTTCTTCTGTA 1260
QY 3854 CTTTCTTCAGTTGTTTCTCTCTTTGCTGCTGTTGCTGCTGTTGTTGGAAGATCAAAC 3913
Db 1261 CTTTCTTCAGTTGTTTCTCTCTTTGCTGCTGTTGCTGCTGTTGTTGGAAGATCAAAC 1320
```

```
QY 3914 AAAGTTGTTGGCCCTCCAGACCTGAGAGCAAACTTTCTTCGAGAGATGCAACAGATGGCCA 3973
Db 1321 AAAGTTGTTGGCCCTCCAGACCTGAGAGCAAACTTTCTTCGAGAGATGCAACAGATGGCCA 1380
QY 3974 GCCGTCCCTTTCCCTCTGTAATGTCGGCTTGGAAACAGATGAGAGCCTCTCTGATCTTA 4033
Db 1381 GCCGTCCCTTTCCCTCTGTAATGTCGGCTTGGAAACAGATGAGAGCCTCTCTGATCTTA 1440
QY 4034 TTGGGGGAGATATAAGACTGTTCCTCCAAACCCATTGCACTGGAGCCGTGTTTGGCAACA 4093
Db 1441 TTGGGGGAGATATAAGACTGTTCCTCCAAACCCATTGCACTGGAGCCGTGTTTGGCAACA 1500
QY 4094 AAGCGCTGCTCTCTCTGTTGAGGCTCCCTCGAGGCTGGTGGCATCCCTCTCTC 4153
Db 1501 AAGCGCTGCTCTCTCTGTTGAGGCTCCCTCGAGGCTGGTGGCATCCCTCTCTC 1560
QY 4154 CTGGGCACTGAGCTCTGCTGTGGCCAGCGCTGGTGGACATTTCTCAGCAGATGCCGA 4213
Db 1561 CTGGGCACTGAGCTCTGCTGTGGCCAGCGCTGGTGGACATTTCTCAGCAGATGCCGA 1620
QY 4214 TAGTGTACAGGAGAGTCAAGAGCCGTGAGAAACCGAAGCAGCAGCCCTGCAAGC 4273
Db 1621 TAGTGTACAGGAGAGTCAAGAGCCGTGAGAAACCGAAGCAGCAGCCCTGCAAGC 1680
QY 4274 CTGGGACCTGCACTGA 4290
Db 1681 CTGGGACCTGCACTGA 1697

RESULT 7
AR164815 LOCUS 6370 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 12 from patent US 6274339.
ACCESSION AR164815
VERSION AR164815.1 GI:16238024
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6370)
AUTHORS Moore,K. and Nagle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body
weight disorders, including obesity
JOURNAL Patent: US 6274339-A 12 14-AUG-2001;
FEATURES Location/Qualifiers
source 1..6370
/organism="unknown"
/moi_type="unassigned DNA"

ORIGIN
Query Match 35.2%; Score 1510; DB 6; Length 6370;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2781 GCAGTCCCGGACACCATGTGCTTGGAGCAGCAGCATGTGGATTCACACGCGGCGCTC 2840
Db 548 GCAGTCCCGGACACCATGTGCTTGGAGCAGCAGCATGTGGATTCACACGCGGCGCTC 607
QY 2841 TGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTTCCCAATGCTATGTGGCTC 2900
Db 608 TGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTTCCCAATGCTATGTGGCTC 667
QY 2901 CTTCCCTTTTGGCCAGTGTATGGAATGATACATGAGCAGCCTGCCCTGAAATTTG 2960
Db 668 CTTCCCTTTTGGCCAGTGTATGGAATGATACATGAGCAGCCTGCCCTGAAATTTG 727
QY 2961 TTCAGGCTACTGTACTCTGTAGTTCATTGCTTGGAGCAACAGGCTGTGGTGTACTGA 3020
Db 728 TTCAGGCTACTGTACTCTGTAGTTCATTGCTTGGAGCAACAGGCTGTGGTGTACTGA 787
QY 3021 TCCAGCAATATCTGCGAAAGGAAATGCAATAGAGGTTCTTATAAAGACAGATGAGAT 3080
Db 788 TCCAGCAATATCTGCGAAAGGAAATGCAATAGAGGTTCTTATAAAGACAGATGAGAT 847
```

```
QY 3081 GCCTTCGCAAGCCCTACAGGAATTTCTATCCACAGCCCTCTGCTCAATTTCCAGCATGTG 3140
Db 848 GCCTTCGCAAGCCCTACAGGAATTTCTATCCACAGCCCTCTGCTCAATTTCCAGCATGTG 907
QY 3141 TCTAGAGGACGAGATACAACTGGCTTTTCATTTCACTGTCAGCTTGCCTCAATTTCCAGCATGTG 3200
Db 908 TCTAGAGGACGAGATACAACTGGCTTTTCATTTCACTGTCAGCTTGCCTCAATTTCCAGCATGTG 967
QY 3201 CCACAGTAAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCAAGGCAA 3260
Db 968 CCACAGTAAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCAAGGCAA 1027
QY 3261 GCCTGCGAGACCTGCATATCTGGCTTCTAGGTGATCCACCAATGAGGGGAAATGTCA 3320
Db 1028 GCCTGCGAGACCTGCATATCTGGCTTCTAGGTGATCCACCAATGAGGGGAAATGTCA 1087
QY 3321 GCATGCAAGTGAATGGGACCGCTCTGTGTCACACCAACACGGGCAAGTGTCTTG 3380
Db 1088 GCATGCAAGTGAATGGGACCGCTCTGTGTCACACCAACACGGGCAAGTGTCTTG 1147
QY 3381 CACCAACAGGGCGTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAATCATACCA 3440
Db 1148 CACCAACAGGGCGTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAATCATACCA 1207
QY 3441 AGGAACCCCTCAGAGGACATGTTATTATCTCTTATTGACTATCAGTTCACCTT 3500
Db 1208 AGGAACCCCTCAGAGGACATGTTATTATCTCTTATTGACTATCAGTTCACCTT 1267
QY 3501 TAGTCTATCCAGGAAGATCGCTATTACACAGCTATCAATTTTGTGGTACTCTCTGA 3560
Db 1268 TAGTCTATCCAGGAAGATCGCTATTACACAGCTATCAATTTTGTGGTACTCTCTGA 1327
QY 3561 CGAACAAACAGGATTTGACATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACAT 3620
Db 1328 CGAACAAACAGGATTTGACATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACAT 1387
QY 3621 CACTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTTGTCTTC 3680
Db 1388 CACTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTTGTCTTC 1447
QY 3681 AAAAAACAATTAGAGATACAAAGATGTTTCTCTAATGAGAGTGTGATTTTCGCAA 3740
Db 1448 AAAAAACAATTAGAGATACAAAGATGTTTCTCTAATGAGAGTGTGATTTTCGCAA 1507
QY 3741 CCACCCAAATACACTTTCTTTGTTTATGTACAGTAAATTCACCTGCCATCAAAATCCA 3800
Db 1508 CCACCCAAATACACTTTCTTTGTTTATGTACAGTAAATTCACCTGCCATCAAAATCCA 1567
QY 3801 GATTGCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTCTTCTGTCATTTCTT 3860
Db 1568 GATTGCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTCTTCTGTCATTTCTT 1627
QY 3861 CAGTGTGTTCTCTCTTTGCTGCTGGTGGTCTGTGGTTTGAAGATCAAAAGATTTG 3920
Db 1628 CAGTGTGTTCTCTCTTTGCTGCTGGTGGTCTGTGGTTTGAAGATCAAAAGATTTG 1687
QY 3921 TTGGGCTCCAGAGCTAGAGACAACTTCTTCAGAGATGCAACAGATGCGCAGCGCTCC 3980
Db 1688 TTGGGCTCCAGAGCTAGAGACAACTTCTTCAGAGATGCAACAGATGCGCAGCGCTCC 1747
QY 3981 CTTTGCTCTGTAATGTGCTGTTGGAACAGATGAGGAGCTCTCTGATCTTATTGGGG 4040
Db 1748 CTTTGCTCTGTAATGTGCTGTTGGAACAGATGAGGAGCTCTCTGATCTTATTGGGG 1807
QY 4041 GAGTATAAGACTGTTCCCAAAACCATTCACATGGAGCGCTGTTTGGCAACAAAGCCG 4100
Db 1808 GAGTATAAGACTGTTCCCAAAACCATTCACATGGAGCGCTGTTTGGCAACAAAGCCG 1867
QY 4101 TGTCTCTCTGTGTTGTGAGGCTCCCTCGAGGCTCGGTGGGATCCCTCTCTCTGGGCA 4160
Db 1868 TGTCTCTCTGTGTTGTGAGGCTCCCTCGAGGCTCGGTGGGATCCCTCTCTCTGGGCA 1927
```

```
QY 4161 GTCAGGTCCTGTTGGCCAGCGCCCTGGTGAGACATTTCTCAGCAGATGCCGATAGTGA 4220
Db 1928 GTCAGGTCCTGTTGGCCAGCGCCCTGGTGAGACATTTCTCAGCAGATGCCGATAGTGA 1987
QY 4221 CAAAGGAGAGTCAGAGCGCTGAGAAACCGGAAGCAGCAGCCCTGCACAGCTGGGAC 4280
Db 1988 CAAAGGAGAGTCAGAGCGCTGAGAAACCGGAAGCAGCAGCCCTGCACAGCTGGGAC 2047
QY 4281 CTGCATCTGA 4290
Db 2048 CTGCATCTGA 2057

RESULT 8
AB011120
LOCUS Homo sapiens mRNA for KIAA0548 protein, partial cds. PRI 10-APR-1998
DEFINITION AB011120
ACCESSION AB011120
VERSION AB011120.1 GI:3043619
KEYWORDS KIAA0548 protein.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Nagase,T., Ishikawa,K., Miyajima,N., Tanaka,A., Kotani,H.,
Nomura,N. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes. IX.
The complete sequences of 100 new cDNA clones from brain which can
code for large proteins in vitro
JOURNAL DNA Res. 5 (1), 31-39 (1998)
MEDLINE 98290545
PUBMED 9628581
REFERENCE 2 (bases 1 to 5632)
AUTHORS Chara,O., Nagase,T. and Ishikawa,K.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute,
DNA Technology; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
(E-mail:cdnaif@kazusa.or.jp, Tel:++81-438-52-3913,
Fax:++81-438-52-3914)
FEATURES
Location/Qualifiers
source 1..5632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HH0769"
/sex="male"
/tissue_type="brain"
/clone_lib="pBluescriptII SK plus"
gene 1..5632
/gene="KIAA0548"
<1..1361
CDS /gene="KIAA0548"
/codon_start=3
/product="KIAA0548 protein"
/protein_id="BAA25474.1"
/db_xref="GI:3043620"
/translation="TMGTCPPEKSGYCTSHCLEQPCGWCCTDPSTNGKCKTEGSK
KGPVMPQSAPTGNFYQPLINSSMCLEDSRYNWSFIHCPCQCNHSGKICINQSIQCEK
CENLTGKHCETCISGYDPTNGKQPCCKNGHASLQNTGKCTCTGKVGKDEC
QLCVENRYQGNPLRGTCYITLLIDYQFTSLSQEDDRYITAINFVATPDQNRDLDM
FINAKNPNLNIITWAASFAGTQAGEMPVVSFKNIKEYKDSFNKFPDFNHNITP
FVYVSNTFWPIKIQAFSQHSNFDLVQFFVTFSCFLSLLVAAVVMKIKQSCWASR
RREQLLRMQMASRPFASVNVALETDEPPDLIGGSIKTVPKETIALEPCFGNKAVAL
SVFVRLPRGLGIPPPQSGSLAVASALVDISQWPIVYVKEKSGAVNRKQPPAPQGT
CI"
```

## ORIGIN

Query Match 31.7%; Score 1361; DB 9; Length 5632;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1361; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2930 ATACGATGAGCACCTGCCCCCTCGAAATTTCTTCAAGGCTACTGTACCTGTAGTCTGCT 2989

Db 1 ATACGATGAGCACCTGCCCCCTGAAATTTGTTCAAGCTACTGTACCTGTAGTCAATGCT 60  
2990 TGGAGCAACAGGCTGTGGCTGTGTACTGATCCAGCATATCTGCAAGGGAATGCA 3049  
Db 61 TGGAGCAACAGGCTGTGGCTGTGTACTGATCCAGCATATCTGCAAGGGAATGCA 120  
3050 TAGAGGGTTCTATAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCT 3109  
Db 121 TAGAGGGTTCTATAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCT 180  
3110 ATCCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGGAGATCAACTGTGCTT 3169  
Db 181 ATCCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGGAGATCAACTGTGCTT 240  
3170 TCATTCAGTGTCCAGCTTGGCAATGCAACGCGCCACAGTAAATGATCATCAATCAGAGCATCT 3229  
Db 241 TCATTCAGTGTCCAGCTTGGCAATGCAACGCGCCACAGTAAATGATCATCAATCAGAGCATCT 300  
3230 GTGAGAGTGTGAGAACCTGACACAGGCAAGCACTGCGAGACTGCGATATCTGGCTTCT 3289  
Db 301 GTGAGAGTGTGAGAACCTGACACAGGCAAGCACTGCGAGACTGCGATATCTGGCTTCT 360  
3290 ACGTGTATCCCAACCAATGGAGGGAATGTGAGCATGCAATGCAATGGGCAAGCTCTC 3349  
Db 361 ACGTGTATCCCAACCAATGGAGGGAATGTGAGCATGCAATGCAATGGGCAAGCTCTC 420  
3350 TGTGCAACCAACCAACGCGCAAGTGTCTGACCAACCAAGGGGCTCAAGGGGAGGAGT 3409  
Db 421 TGTGCAACCAACCAACGCGCAAGTGTCTGACCAACCAAGGGGCTCAAGGGGAGGAGT 480  
3410 GCCAGCTATGTGAGGTGAGAAATGCGATACAGAGAAACCTCTCAGAGGAACATGTATT 3469  
Db 481 GCCAGCTATGTGAGGTGAGAAATGCGATACAGAGAAACCTCTCAGAGGAACATGTATT 540  
3470 ATACTCTTCTTATGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTATT 3529  
Db 541 ATACTCTTCTTATGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTATT 600  
3530 ACACAGCTATCAATTTTGTGCTACTCTGACGAAACAAACAGGGAATTTGGAATGTTCA 3589  
Db 601 ACACAGCTATCAATTTTGTGCTACTCTGACGAAACAAACAGGGAATTTGGAATGTTCA 660  
3590 TCAATGCTTCAAGAAATTCACCTCAACATCACTGCGGTGCGAGTTTCTCAGCTGGAA 3649  
Db 661 TCAATGCTTCAAGAAATTCACCTCAACATCACTGCGGTGCGAGTTTCTCAGCTGGAA 720  
3650 CCCAGGCTGGAGAGAGATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATA 3709  
Db 721 CCCAGGCTGGAGAGAGATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATA 780  
3710 GTTCTCTAATGAGAGTTTGAATTTTGGCAACCCCAATATACATTTCTTTGTTATG 3769  
Db 781 GTTCTCTAATGAGAGTTTGAATTTTGGCAACCCCAATATACATTTCTTTGTTATG 840  
3770 TCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCTTCTCTCAGCAGCAATTTTA 3829  
Db 841 TCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCTTCTCTCAGCAGCAATTTTA 900  
3830 TGGACCTGGATACAGTTCTTGTGATCTTTCTTCAAGTTGTTCTCTCTTCTCTGCTG 3889  
Db 901 TGGACCTGGATACAGTTCTTGTGATCTTTCTTCAAGTTGTTCTCTCTTCTCTGCTG 960  
3890 CTGCTGTGTTTGAAGATCAACAAAGTTGTTTGGGCTCCAGACGCTAGAGAGCACTTC 3949  
Db 961 CTGCTGTGTTTGAAGATCAACAAAGTTGTTTGGGCTCCAGACGCTAGAGAGCACTTC 1020  
3950 TTCAGAGATCAACAGATGCCAGCGCTCCCTTTGCTCTGTAAATGTGCTTGGAA 4009  
Db 1021 TTCAGAGATCAACAGATGCCAGCGCTCCCTTTGCTCTGTAAATGTGCTTGGAA 1080  
4010 CAGATGAGGACCTCTGATCTTATTTGGGGGAGTATAAAGACTGTTCCTCCAAACCCATG 4069

Db 1081 CAGATGAGGACCTCTCTGATCTTATTTGGGGAGATATAAGACTGTTCCCAACCCATTG 1140  
4070 CACTGGAGCCGTGTTTGGCAACAAAGCCGCTGCTCTCTGTTGTTGAGGCTCCCTC 4129  
Db 1141 CACTGGAGCCGTGTTTGGCAACAAAGCCGCTGCTCTCTGTTGTTGAGGCTCCCTC 1200  
4130 GAGGCTGGGTGGCATCCCTCTCTGCGGCAAGTCAAGTCTGCTGTCGCGAGCCCTGG 4189  
Db 1201 GAGGCTGGGTGGCATCCCTCTCTGCGGCAAGTCAAGTCTGCTGTCGCGAGCCCTGG 1260  
4190 TGGACATTTCTCAGCAGATGCCGATAGTGTACAAAGGAGAGTCAAGGACCGTGAGAAAC 4249  
Db 1261 TGGACATTTCTCAGCAGATGCCGATAGTGTACAAAGGAGAGTCAAGGACCGTGAGAAAC 1320  
4250 GGAAGCAGACGCCCTCTGACAGCTGGGACCTGCATCTGA 4290  
Db 1321 GGAAGCAGACGCCCTCTGACAGCTGGGACCTGCATCTGA 1361

RESULT 9  
AX778545  
LOCUS  
DEFINITION Sequence 702 from Patent WO03039443.  
ACCESSION AX778545  
VERSION AX778545.1 GI:32695539  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Haferlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,  
Duges, M., Ellis, R., Brors, B. and Mergenthaler, S.  
TITLE Novel genetic markers for leukemias  
JOURNAL Patent: WO 03039443-A 702 15-MAY-2003;  
Deutsches Krebsforschungszentrum (DE);  
Ludwig-Maximilian-Universitaet Muenchen (DE);  
PD Dr. Dr. (DE); Schoch, Claudia (DB); Kern, Wolfgang (DE)

FEATURES  
Location/Qualifiers  
1..543  
/organism="Homo sapiens"  
/mol\_type="unasigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 12.7%; Score 543; DB 6; Length 543;  
Best/Local Similarity 100.0%; Pred. No. 2.5e-281; Indels 0; Gaps 0;  
Matches 543; Conservative 0; Mismatches 0

QY 3250 ACCACAGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCAATGGA 3309  
Db 1 ACCACAGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCAATGGA 60  
QY 3310 GGGAAATGTCAAGCAATGCAATGGCAACCGCTCTGTGCAACACCAACACACGCGC 3369  
Db 61 GGGAAATGTCAAGCAATGCAATGGCAACCGCTCTGTGCAACACCAACACACGCGC 120  
QY 3370 AAGTGTCTTCTGCAACCAACCAAGGGCTCAAGGGGGACGAGTCCAGCTATGTAGGTAGAA 3429  
Db 121 AAGTGTCTTCTGCAACCAACCAAGGGCTCAAGGGGGACGAGTCCAGCTATGTAGGTAGAA 180  
QY 3430 AATGATACCAAGGAACCTCTCAGAGGAACATGTTATTTACTCTCTTTATTGACTAT 3489  
Db 181 AATGATACCAAGGAACCTCTCAGAGGAACATGTTATTTACTCTCTTTATTGACTAT 240  
QY 3490 CAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACAGAGTATCAATTTTGTG 3549  
Db 241 CAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACAGAGTATCAATTTTGTG 300  
QY 3550 GCTACTCTGAGCAACAAACAGGAGTTTGACATGTTTCATCAATGCTTCCAGAAATTC 3609  
Db 301 GCTACTCTGAGCAACAAACAGGAGTTTGACATGTTTCATCAATGCTTCCAGAAATTC 360







```

Db      48926 CCCTGTGTACACGCGGTGCTGCAACCTTGGACCGCGAGTGGTCTGCCCGCGGC 48985
Oy      370  TGGTGGCGGAGCAATGCGACCATGCGGGGGCGGCTTCAG 410
Db      48986 TGGTGGCGGAGCAATGCGACCATGCGGGGGCGGCTTCAG 49026

RESULT 12
AC023420 Homo sapiens chromosome 2 clone RP11-737G10 map 2, WORKING DRAFT
LOCUS     AC023420 155869 bp DNA linear HTG 23-SEP-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-737G10 map 2, WORKING DRAFT
SEQUENCE, 15 unordered pieces.
ACCESSION AC023420
VERSION    AC023420.4 GI:10280876
KEYWORDS  HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 155869)
AUTHORS  Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavskiy,L.,
Bouckhalter,B., Brown,A., Burkett,G., Campolano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
Dearellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
Fenster,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
Garagan,J., Gardyna,S., Gande,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Landers,T., Larcocque,K., Lehoczyk,J., Levine,R.,
Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M.,
McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Minova,T., Miranda,C., Mienga,V., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M.,
Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Subramanian,A., Talamas,J., Testave,S., Theodore,J., Tirrell,A.,
Travers,M., Triggilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
Zody,M.
Direct Submission
Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 23, 2000 this sequence version replaced gi:7767820.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L6805
Center clone name: 737 G.10
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 145736 bases at least Q40
Consensus quality: 151534 bases at least Q30
Consensus quality: 153551 bases at least Q20
Insert size: 157000; agarose-fp
Quality coverage: 4.2 in Q20 bases; agarose-fp
Quality coverage: 4.3 in Q20 bases; sum-of-contigs
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 12284: contig of 12284 bp in length
* 12285 12384: gap of 100 bp
* 12385 13704: contig of 1320 bp in length
* 13705 15657: contig of 1853 bp in length
* 15658 17557: gap of 100 bp
* 17558 17755: contig of 1998 bp in length
* 17756 17855: gap of 100 bp
* 17856 19875: contig of 2020 bp in length
* 19876 22894: contig of 2919 bp in length
* 22895 26992: contig of 3998 bp in length
* 26993 27092: gap of 100 bp
* 27093 31422: contig of 4330 bp in length
* 31423 35220: contig of 6398 bp in length
* 35221 38020: gap of 100 bp
* 38021 48644: contig of 10624 bp in length
* 48645 48744: gap of 100 bp
* 48745 69112: contig of 20368 bp in length
* 69113 86049: contig of 16837 bp in length
* 86050 86150: gap of 100 bp
* 86151 107477: contig of 21328 bp in length
* 107478 131094: contig of 23517 bp in length
* 131095 131194: gap of 100 bp
* 131195 155869: contig of 24675 bp in length.
*
* Location/Qualifiers
*   1..155869
*     /organism="Homo sapiens"
*     /mol_type="genomic DNA"
*     /db_xref="taxon:9606"
*     /chromosome="2"
*     /map="2"
*     /clone="RP11-737G10"
*     /clone_lib="RPC1-11 Human Male BAC"
*   1..12284
*     /note="assembly_fragment"
*     /clone_end:SP6
*     /vector_side:left"
* 12385..13704
*     /note="assembly_fragment"
* 13805..15657
*     /note="assembly_fragment"
* 15758..17755
*     /note="assembly_fragment"
* 17856..19875
*     /note="assembly_fragment"
* 19976..22894
*     /note="assembly_fragment"
* 22995..26992
*     /note="assembly_fragment"
* 27093..31422
*     /note="assembly_fragment"
* 31523..37920
*     /note="assembly_fragment"
* 38021..48644
*     /note="assembly_fragment"
* 48745..69112
*     /note="assembly_fragment"
* 69213..86049
*     /note="assembly_fragment"
* 86150..107477

```

## FEATURES

## source

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

```

/note="assembly_fragment"
misc_feature
107578..131094
/note="assembly_fragment"
131195..155869
/note="assembly_fragment"

```

```
Query Match      7.0%; Score 299; DB 2; Length 155869;
Best Local Similarity 99.5%; Pred. No. 1.5e-149;
Matches 399; Conservative 0; Mismatches 2; Indels 0;
```

10	GCAGCGCGGCACCTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGCGAGCGCTCGCG	59
142651	GACGCGCGCGGCACCTGAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGCGAGCGCTCGCG	142710
70	GCGAGGACGGCGGCGCGACCTGGGACTGGGACGTGACACAGGGCTGGGAGGCCGGGGCTG	129
142711	GCGAGGACGGCGGCGCGACCTGGGACTGGGACGTGACACAGGGCTGGGAGGCCGGGGCTG	142770
130	GGGGCGGGGCTGCGGCTCCCCGGGCTGTGTCTCCACCGTGGGGCCACGGCTGTGCTG	189
142771	GGGGCGGGGCTGCGGCTCCCCGGGCTGTGTCTCCACCGTGGGGCCACGGCTGTGCTG	142830
190	GTGCTGTGTGTCTCCGCGCGCGCGCTGTGCTGTGCTGTGCTGTGAGGCCGAGGGC	249
142831	GTGCTGTGTGTCTCTCGCGCGCGCTGTGCTGTGCTGTGCTGTGAGGCCGAGGGC	142890
250	GCGGCGGGCGGCGGCGGGTGTGCGGCTCAGCCGCGACGCGAGGCCAAGGAATGTGACCG	309
142891	GCGGCGGGCGGCGGCGGGTGTGCGGCTCAGCCGCGACGCGAGGCCAAGGAATGTGACCG	142950
310	CCCTGTGTCAACGGCGGTGCTGCAACCCCTGGCACCGGCGAGTGCCTGCCCGCGCGGC	369
142951	CCCTGTGTCAACGGCGGTGCTGCAACCCCTGGCACCGGCGAGTGCCTGCCCGCGCGGC	143010
370	TGGGTGGCGAGCAATGCGACACTGCGGGGGCGCTTCAG	410
143011	TGGGTGGCGAGCAATGCGACACTGCGGGGGCGCTTCAG	143051

[illegible]

Unclassified.  
1 (bases 1 to 8827)  
Moore,K. and Nagle,D.Lynn.  
METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT OF BODY  
TITLE weight disorders, including obesity  
Patent: US 6274339-A 1 14-AUG-2001;  
JOURNAL Location/Qualifiers  
FEATURES 1..8827  
source

	ORIGIN	/mol_type= unassigned dna
	Query Match	6.8%; Score 291; DB 6; Length 8827;
	Best Local Similarity	100.0%; Pred.No. 4.6e-145;
	Matches 291; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	4000 GCCTTGGAAAACAGATGAGGAGCCTCCTGTATTATTGGGGGAGTATAAAGACTGTCCC	4059
Db	4266 GCCTTGGAAAACAGATGAGGAGCCTCCTGTATTATTGGGGGAGTATAAAGACTGTCCC	4325
QY	4060 AAACCCATTGCACATGGAGCGTGTTTGGCAACAAGCCGTCCTCTCTGTGTTTTGTG	4119
Db	4326 AATCCCATTTGTAATCTGGAGCGTGTTTTTGACAAAGAAGCGCTGCTCTCTGTGTTTTGTG	4385

QY	4120	AGGCTCCCTCGAGAGCCTGGTGGCATCCCTCTCTCTGGCAGTCAGGTCCTGCTGGGCC	41179
DB	4386	AGGCTCCCTCGAGAGCCTGGTGGCATCCCTCTCTCTGGCAGTCAGGTCCTGCTGGGCC	4445
QY	4180	AGCGCCCTGGTGGCATTCTTCTCAGCAGATGCCGATAGTGTA CAAGAGAAGTCAGAGGCC	4239
DB	4446	AGCGCCCTGGTGGCATTCTTCTCAGCAGATGCCGATAGTGTA CAAGAGAAGTCAGAGGCC	4505
QY	4240	GTGAGAAACCGGAAGCAGCAGAGCCCTCTGCACAGCCTGGGACCTGCATCTGA	4290
DB	4506	GTGAGAAACCGGAAGCAGCAGAGCCCTCTGCACAGCCTGGGACCTGCATCTGA	4556

RESULT 14  
 AR273223/c  
 LOCUS 1  
 DEFINITION  
 ACCESSION AR273223  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 FEATURES  
 source  
 ORIGIN

Query Match	5.7%;	Score 246;	DB 6;	Length 246;
Best Local Similarity	100.0%;	Fred. No. 1.6e-120;		
Matches 246;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2515	CTCGGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC	2574	
Db	246	CTCGGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTCGGC	187	
QY	2575	CTTCGGAGACATCAATGTGTCCTACTGGTCTCGGAAGATATGTCGCCATTTTACAAATAGT	2634	
Db	186	CTTCGGAGACATCAATGTGTCCTACTGGTCTCGGAAGATATGTCGCCATTTTACAAATAGT	127	
QY	2635	TTACTACAGTGGATCCGTCCTGAGCCACGTGATCTGTGGATTTCTGTGGAAATTTTATCAGAA	2694	
Db	126	TTACTACAGTGGATCCGTCCTGAGCCACGTGATCTGTGGATTTCTGTGGAAATTTTATCAGAA	67	
QY	2695	CCCACTACTCGGGGACTGAAGCGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGTCTGT	2754	
Db	66	CCCACTACTCGGGGACTGAAGCGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGTCTGT	7	
QY	2755	GAAAGG	2760	
Db	6	GAAAGG	1	

RESULT	15
AR276804/c	
LOCUS	AR276804
DEFINITION	Sequence 956 from patent US 6509448.
ACCESSION	AR276804
VERSION	AR276804.1 GI:29710451
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
	Unclassified.
REFERENCE	1 (bases 1 to 246)
AUTHORS	Wang,T., Bangur,C.S., Lodes,M.J., Fanger,G.R., Vedvick,T.S., Carter,D., Retter,M.W., Mannion,J., Fan,L. and Wang,A.
	linear PAT 10-APR-2003

TITLE Compositions and methods for the therapy and diagnosis of lung cancer  
JOURNAL Patent: US 6509448-A 966 21-JAN-2003;  
FEATURES Location/Qualifiers  
source 1. .246  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 5.7%; Score 246; DB 6; Length 246;  
Best Local Similarity 100.0%; Pred. No. 1.6e-120;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATATGAGTCTCAGAGCATGTCACAGCTCAAGCTCACTTAACCCCATGGTGGC 2574  
Db 246 CTGCGAATATGAGTCTCAGAGCATGTCACAGCTCAAGCTCACTTAACCCCATGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTGCTCTACTGCTGCTGGAAGATATGTCCTCAATTTACAAATAGT 2634  
Db 186 CTTGCGAAGATCAATGTGCTCTACTGCTGCTGGAAGATATGTCCTCAATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCGGTCTGAGCCAGTGGATCTGCTGGAATTTATCAGAA 2694  
Db 126 TTACTACAGTGGATGCGGTCTGAGCCAGTGGATCTGCTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 2754  
Db 66 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 7

QY 2755 GAAAGG 2760  
Db 6 GAAAGG 1

RESULT 16  
AR407079/c  
LOCUS AR407079 246 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 966 from patent US 6630574.  
ACCESSION AR407079  
VERSION AR407079.1 GI:40156890  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 246)  
AUTHORS Wang, T., Bangur, C.S., Lodes, M.J., Fanger, G.R., Vedvick, T.S., Carter, D., Retter, M.W., Mannion, J. and Ran, L.  
TITLE Compositions and methods for the therapy and diagnosis of lung cancer  
JOURNAL Patent: US 6630574-A 966 07-OCT-2003;  
FEATURES Location/Qualifiers  
source 1. .246  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 5.7%; Score 246; DB 6; Length 246;  
Best Local Similarity 100.0%; Pred. No. 1.6e-120;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATATGAGTCTCAGAGCATGTCACAGCTCAAGCTCACTTAACCCCATGGTGGC 2574  
Db 246 CTGCGAATATGAGTCTCAGAGCATGTCACAGCTCAAGCTCACTTAACCCCATGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTGCTCTACTGCTGCTGGAAGATATGTCCTCAATTTACAAATAGT 2634  
Db 186 CTTGCGAAGATCAATGTGCTCTACTGCTGCTGGAAGATATGTCCTCAATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCGGTCTGAGCCAGTGGATCTGCTGGAATTTATCAGAA 2694  
Db 126 TTACTACAGTGGATGCGGTCTGAGCCAGTGGATCTGCTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 2754

TITLE Compositions and methods for the therapy and diagnosis of lung cancer  
JOURNAL Patent: US 6509448-A 966 21-JAN-2003;  
FEATURES Location/Qualifiers  
source 1. .246  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 5.7%; Score 246; DB 6; Length 246;  
Best Local Similarity 100.0%; Pred. No. 1.6e-120;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATATGAGTCTCAGAGCATGTCACAGCTCAAGCTCACTTAACCCCATGGTGGC 2574  
Db 246 CTGCGAATATGAGTCTCAGAGCATGTCACAGCTCAAGCTCACTTAACCCCATGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTGCTCTACTGCTGCTGGAAGATATGTCCTCAATTTACAAATAGT 2634  
Db 186 CTTGCGAAGATCAATGTGCTCTACTGCTGCTGGAAGATATGTCCTCAATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCGGTCTGAGCCAGTGGATCTGCTGGAATTTATCAGAA 2694  
Db 126 TTACTACAGTGGATGCGGTCTGAGCCAGTGGATCTGCTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 2754

Db 66 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 7  
QY 2755 GAAAGG 2760  
Db 6 GAAAGG 1

RESULT 17  
AX368256/c  
LOCUS AX368256 246 bp DNA linear PAT 16-FEB-2002  
DEFINITION Sequence 966 from Patent WO0204514.  
ACCESSION AX368256  
VERSION AX368256.1 GI:18856329  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1  
AUTHORS Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W., Marnerakis, M., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S., McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.  
TITLE Compositions and methods for the therapy and diagnosis of lung cancer  
JOURNAL Patent: WO 0204514-A 966 17-JAN-2002;  
FEATURES Location/Qualifiers  
source 1. .246  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 5.7%; Score 246; DB 6; Length 246;  
Best Local Similarity 100.0%; Pred. No. 1.6e-120;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATATGAGTCTCAGAGCATGTCACAGCTCAAGCTCACTTAACCCCATGGTGGC 2574  
Db 246 CTGCGAATATGAGTCTCAGAGCATGTCACAGCTCAAGCTCACTTAACCCCATGGTGGC 187

QY 2575 CTTGCGAAGATCAATGTGCTCTACTGCTGCTGGAAGATATGTCCTCAATTTACAAATAGT 2634  
Db 186 CTTGCGAAGATCAATGTGCTCTACTGCTGCTGGAAGATATGTCCTCAATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCGGTCTGAGCCAGTGGATCTGCTGGAATTTATCAGAA 2694  
Db 126 TTACTACAGTGGATGCGGTCTGAGCCAGTGGATCTGCTGGAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 2754  
Db 66 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 7

QY 2755 GAAAGG 2760  
Db 6 GAAAGG 1

RESULT 18  
H009384S15  
LOCUS H009384S15 3257 bp DNA linear PRI 14-DEC-2000  
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exons 17 and 18.  
ACCESSION AF218903  
VERSION AF218903.1 GI:8118069  
KEYWORDS 15 of 27  
SEGMENT Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 3257)  
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.



```

DVTQMTWILKDSRFRYLHTAVIVSGTLMVFGNTHNDTSMHGAKCFSSDFMAYDI
ACDRWSLPRPLDHHVDFGHSVAVLHNSMTVFGFNSLLSDILVFTSEOCDAHRS
ACDLAAGVGRGCVNWTGSGOCSIALWALDEOEKLSKFCSEKRLDHRCDQHTDCY
STANTNDCHWCHDCHVBNHSCSEGOISIFREYENCPCDPMYCNKXTSCRSALDQ
NQCEPRNOCETALPENICIGIWHLVGNSCLIKTTAKENDKLCFNRHNSALLSLT
TQKVEFVLQJRMGSSOSMSKLLTTPWGLRKINYSYCWEDMSPFTNSLLOWPS
ESDAGCGILSEPSTRGKAATCINPLNGSVCPERPAHSAKQCRTPCALRTACGDCT
SSGSCMWCNQCVDNSNAYVASFPFGQMEWYTMSTPCPCSCGYCTCSHCLIQPG
CWCQTDPSNTGKICIEGYSKVPKMPQAPTFYPPOLNNSMCLSDRSRWNSFIH
CPACQCNHSGKINOSICEKCNLTGKHCTCISGFYDPTNGKCKPCCKNGHASL
CNNTGKCFCTYGVKGDSCOLCEVENRYQGNPLSGTCYVTLIDYQTFSLSQBDDR
YXTALNFVATPEQNDLDMFINASKNFNLNTWASFSAGTQAGZEMPVVSKNIKE
YKDSFSEKEDFRNPNITTFYVNSFTWPKIQIAFSOHSNFMDLVFFVTFSCFL
SLLVAUVWKKIQCSWASREQLREMQQWASRPASVNAVLEDEPDLIGSSI
KITVPKIALPECFGNKAALVSVLRPLGLGIPPGGSLVASALVDISQOMPIVY
KEKSGAVNRKQPPAQPGCTCI"
join(<2. .161.1842. .2010. 12065. 12155.13334. .13577.
14948. 15131.15881. 16035.18020. .18104.25403. .25623,
28499. .28620.29531. .29669.31254. .31444.34807. .35026,
35570. .36755.37319. .37552.43841. .43978.47151. .47244,
49049. .49101.50578. .50672.53618. .53696.56777. .56934,
68500. .68517)
/gene="ATRN"
/note="match: proteins: Tr:O75882 Tr:Q9NZ58 Tr:O95414"
/codon_start=1
/evidence=not experimental
/product="G074LH3.1.1 (attractin (with dipeptidylpeptidase
IV activity) secreted isoform)"
/protein_id="CAC32456.1"
/db_xref="GI:13160051"
/db_xref="GOA:O75882"
/db_xref="SWISS-PROT:O75882"
/translation="ISNSSDVFCECSNWEKGACDIPHTDNCGPFHRGICNSDVR
QCSFSDQWQFGCVFVFNANQSVFTREIYSNLKUPRASHKAVNGNINWVGGYMFNE
SDYNMLVLAIDRFLWPLNRSVNNVYVYGHSLAYKDKIYVIGGKIDSTGNTNEL
RVFHTNDSWLLPFAKEQYAVVGHSAHIVLKNGRVMLVI FGHCPLYGIYGNVBE
YDLKNTWSILHTQALVGGYGVSVVDTRALYVHGKAFSAKRYLRADLILRY
DVTQMTWILKDSRFRYLHTAVIVSGTLMVFGNTHNDTSMHGAKCFSSDFMAYDI
ACDRWSLPRPLDHHVDFGHSVAVLHNSMTVFGFNSLLSDILVFTSEOCDAHRS
EACDLAAGVGRGCVNWTGSGOCSIALWALDEOEKLSKFCSEKRLDHRCDQHTDCY
SCTANTNDCHWCHDCHVBNHSCSEGOISIFREYENCPCDPMYCNKXTSCRSALDQ
NQCEPRNOCETALPENICIGIWHLVGNSCLIKTTAKENDKLCFNRHNSALLSLT
TQKVEFVLQJRMGSSOSMSKLLTTPWGLRKINYSYCWEDMSPFTNSLLOWPS
BPSDAGFCGILSEPSTRGKAATCINPLNGSVCPERPAHSAKQCRTPCALRTACGDCT
SSGSCMWCNQCVDNSNAYVASFPFGQMEWYTMSTPCPCSCGYCTCSHCLIQPG
CGWCTDPSNTGKICIEGYSKVPKMPQAPTFYPPOLNNSMCLSDRSRWNSFIH
CPACQCNHSGKINOSICEKCNLTGKHCTCISGFYDPTNGKCKPCCKNGHASL
CNNTGKCFCTYGVKGDSCOLCEVENRYQGNPLSGTCYVTLIDYQTFSLSQBDDR
YXTALNFVATPEQNDLDMFINASKNFNLNTWASFSAGTQAGZEMPVVSKNIKE
YKDSFSEKEDFRNPNITTFYVNSFTWPKIQIAFSOHSNFMDLVFFVTFSCFL
SLLVAUVWKKIQCSWASREQLREMQQWASRPASVNAVLEDEPDLIGSSI
KITVPKIALPECFGNKAALVSVLRPLGLGIPPGGSLVASALVDISQOMPIVY
YKDSFSEKEDFRNPNITTFYVNSFTWPKIQIAFSOHSNFMDLVFFVTFSCFL
17. .569
/gene="ATRN"
/note="match: STS: Em:AL110053 Em:HS741H3S"
903. .990
/note="MER2 repeat: matches 257. .337 of consensus"
991. .1297
/note="AluX repeat: matches 1. .307 of consensus"
1298. .1520
/note="MER2 repeat: matches 1. .257 of consensus"
2718. .2942
/note="AluX repeat: matches 1. .225 of consensus"
2943. .3127
/note="L1P repeat: matches 4678. .4858 of consensus"
complement(3203. .3804)
/note="match: GSS: Em:AQ381369"
3560. .3672
/note="L12 repeat: matches 216. .330 of consensus"
3881. .3977
/note="MER81 repeat: matches 2. .110 of consensus"
4326. .4513
/note="L12 repeat: matches 738. .924 of consensus"
4539. .4682
/note="MER3 repeat: matches 1. .149 of consensus"

```

```

repeat_region 4683. .5195
/note="L12 repeat: matches 1383. .1936 of consensus"
5209. .5762
/note="MLR1E repeat: matches 2. .562 of consensus"
5770. .5835
/note="L12 repeat: matches 1949. .2024 of consensus"
6074. .6708
/note="L12 repeat: matches 2025. .2732 of consensus"
5998. .7305
/note="L12 repeat: matches 1. .309 of consensus"
7663. .7974
/note="AluY repeat: matches 1. .311 of consensus"
complement(9805. .10470)
/note="match: GSS: Em:AQ484998"
9825. .10261
/gene="ATRN"
/note="match: GSS: Em:AQ338818"
9963. .10167
/note="HAL1 repeat: matches 618. .819 of consensus"
10173. .10450
/note="match: GSS: Em:AQ413104"
10902. .11155
/note="L1ME3 repeat: matches 5637. .5902 of consensus"
12706. .12802
/note="L1MB2 repeat: matches 6075. .6168 of consensus"
12803. .13235
/note="L1MA4A repeat: matches 5584. .6168 of consensus"
13713. .14023
/note="AluX repeat: matches 3. .311 of consensus"
14407. .14697
/note="AluJ repeat: matches 1. .293 of consensus"

Query Match 5.5%; Score 234; DB 9; Length 104907;
Best Local Similarity 100.0%; Pred. No. 2.5e-114;
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2951 CTGAAATTTCTCAGGCTACTGTAAGTGTAGTCTGTTGGAGCAACAGGCTGGCT 3010
DB 37319 CTGAAATTTCTCAGGCTACTGTAAGTGTAGTCTGTTGGAGCAACAGGCTGGCT 37378
QY 3011 GGTGTACTGATCCAGCAATCTGCAAGGGAATGATAGAGGTTCTTATAAGGAC 3070
DB 37379 GGTGTACTGATCCAGCAATCTGCAAGGGAATGATAGAGGTTCTTATAAGGAC 37438
QY 3071 CAGTGAAGTGTCTTCCAGCCCTTACAGGAATTTCTTACACAGCCCTCTCAATT 3130
DB 37439 CAGTGAAGTGTCTTCCAGCCCTTACAGGAATTTCTTACACAGCCCTCTCAATT 37498
QY 3131 CCAGCATGTCTTAGAGGACGACAGATACACTGCTTTTCACTTCACTGTCAG 3184
DB 37499 CCAGCATGTCTTAGAGGACGACAGATACACTGCTTTTCACTTCACTGTCAG 37552

RESULT 20
ACOL13324/c
LOCUS ACOL13324
DEFINITION Homo sapiens chromosome 2 clone RP11-388K24 map 2, LOW-PASS
SEQUENCE SAMPLING.
ACCESSION ACOL13324
VERSION ACOL13324.3 GI:9123914
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 118540)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 2, clone RP11-388K24
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 118540)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhgalter,B.,

```



Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,  
 Cooke, P., DeArelano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,  
 Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Horton, L.,  
 Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heatford, A., Klein, J.,  
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Marquis, N.,  
 Lehoczy, J., Lieu, C., Locke, K., McKernan, K., McLaughlin, J., Meldrim, J.,  
 McEwan, P., McDuck, A., Norman, C. H., O'Connor, T., O'Donnell, P.,  
 Morrow, J., Naylor, J., Naylor, R., Roy, A., Santos, R., Severy, P.,  
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
 Testaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

# TITLE JOURNAL

## COMMENT

Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 13, 2000 this sequence version replaced gi:6289165.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L3787

Center clone name: 388\_K\_24

-----

\* NOTE: This record contains 125 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1  
 \* 857 856: contig of 856 bp in length  
 \* 857 956: gap of 100 bp  
 \* 957 1821: contig of 865 bp in length  
 \* 1822 1821: gap of 100 bp  
 \* 1822 2776: contig of 855 bp in length  
 \* 2777 2876: gap of 100 bp  
 \* 2877 3767: contig of 891 bp in length  
 \* 3768 3867: gap of 100 bp  
 \* 3868 4709: contig of 842 bp in length  
 \* 4710 4809: gap of 100 bp  
 \* 4810 5655: contig of 846 bp in length  
 \* 5656 5755: gap of 100 bp  
 \* 5756 6627: contig of 872 bp in length  
 \* 6628 6727: gap of 100 bp  
 \* 6728 7586: contig of 859 bp in length  
 \* 7587 7686: gap of 100 bp  
 \* 7687 8241: contig of 855 bp in length  
 \* 8242 8641: gap of 100 bp  
 \* 8642 9524: contig of 883 bp in length  
 \* 9525 9624: gap of 100 bp  
 \* 9625 10509: contig of 885 bp in length  
 \* 10510 10610: contig of 865 bp in length  
 \* 10610 11474: contig of 865 bp in length  
 \* 11475 11574: gap of 100 bp  
 \* 11575 12441: contig of 867 bp in length  
 \* 12442 12541: gap of 100 bp  
 \* 12542 13409: contig of 868 bp in length  
 \* 13410 13509: gap of 100 bp  
 \* 13510 14439: contig of 930 bp in length  
 \* 14440 14539: gap of 100 bp  
 \* 14540 15433: contig of 894 bp in length  
 \* 15434 15533: gap of 100 bp  
 \* 15534 16442: contig of 909 bp in length  
 \* 16443 16542: gap of 100 bp

16543 17421: contig of 879 bp in length  
 17422 17521: gap of 100 bp  
 17522 18380: contig of 859 bp in length  
 18381 18480: gap of 100 bp  
 18481 19345: contig of 865 bp in length  
 19346 19445: gap of 100 bp  
 20327: contig of 882 bp in length  
 20328 20427: gap of 100 bp  
 21297: contig of 870 bp in length  
 21298 21397: gap of 100 bp  
 22055: contig of 658 bp in length  
 22056 22155: gap of 100 bp  
 22156 23005: contig of 850 bp in length  
 23006 23105: gap of 100 bp  
 24016: contig of 911 bp in length  
 24117 24116: gap of 100 bp  
 24972: contig of 856 bp in length  
 25072: gap of 100 bp  
 25073 25977: contig of 905 bp in length  
 26077: gap of 100 bp  
 26078 26952: contig of 875 bp in length  
 26953 27311: contig of 859 bp in length  
 28011: gap of 100 bp  
 28868: contig of 857 bp in length  
 28968: gap of 100 bp  
 29624: contig of 656 bp in length  
 29724: gap of 100 bp  
 29725 30574: contig of 850 bp in length  
 30575 30674: gap of 100 bp  
 30675 31530: contig of 856 bp in length  
 31531 31630: gap of 100 bp  
 32491: contig of 861 bp in length  
 32492 32591: gap of 100 bp  
 32592 33467: contig of 876 bp in length  
 33468 33567: gap of 100 bp  
 33568 34437: contig of 870 bp in length  
 34438 34537: gap of 100 bp  
 34538 35404: contig of 867 bp in length  
 35405 35504: gap of 100 bp  
 35505 36368: contig of 864 bp in length  
 36369 36468: gap of 100 bp  
 36469 37353: contig of 885 bp in length  
 37354 37453: gap of 100 bp  
 37454 38310: contig of 857 bp in length  
 38311 38410: gap of 100 bp  
 38411 39267: contig of 857 bp in length  
 39268 39367: gap of 100 bp  
 39368 40217: contig of 850 bp in length  
 40218 40317: gap of 100 bp  
 40318 41226: contig of 909 bp in length  
 41227 41326: gap of 100 bp  
 41327 42195: contig of 869 bp in length  
 42196 42295: gap of 100 bp  
 42296 43180: contig of 885 bp in length  
 43181 43280: gap of 100 bp  
 43281 44143: contig of 863 bp in length  
 44144 44244: gap of 100 bp  
 44244 45132: contig of 888 bp in length  
 45133 45231: gap of 100 bp  
 45232 46149: contig of 918 bp in length  
 46150 46249: gap of 100 bp  
 46250 47122: contig of 873 bp in length  
 47123 47222: gap of 100 bp  
 47223 48101: contig of 879 bp in length  
 48102 48201: gap of 100 bp  
 48202 49041: contig of 840 bp in length  
 49042 49141: gap of 100 bp  
 50059: contig of 918 bp in length  
 50159: gap of 100 bp  
 50060 51029: contig of 870 bp in length  
 51030 51129: gap of 100 bp  
 51130 51995: contig of 866 bp in length



Quality coverage: 4.5 in Q20 bases; agarose-fp  
 Quality coverage: 4.6 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 21 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence.  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1 1216: contig of 1216 bp in length
* 1217 1316: gap of 100 bp
* 1317 2985: contig of 1669 bp in length
* 2986 3085: gap of 100 bp
* 3086 4628: contig of 1543 bp in length
* 4629 4729: gap of 100 bp
* 4729 7508: contig of 2779 bp in length
* 7508 7607: gap of 100 bp
* 7607 11292: contig of 3685 bp in length
* 11293 14213: gap of 100 bp
* 14213 14313: contig of 2821 bp in length
* 14313 17828: contig of 3515 bp in length
* 17829 21068: gap of 100 bp
* 21069 21168: contig of 3140 bp in length
* 21169 25035: contig of 3857 bp in length
* 25036 28331: gap of 100 bp
* 28331 33595: contig of 3156 bp in length
* 33596 33695: gap of 100 bp
* 33696 40346: contig of 5164 bp in length
* 40347 40446: gap of 100 bp
* 40447 43807: contig of 6651 bp in length
* 43808 43907: gap of 100 bp
* 43908 48727: contig of 3361 bp in length
* 48728 48827: gap of 100 bp
* 48828 56083: contig of 4820 bp in length
* 56084 56183: gap of 100 bp
* 56184 67055: contig of 7256 bp in length
* 67056 67155: contig of 10882 bp in length
* 67156 82881: gap of 100 bp
* 82882 82981: contig of 15716 bp in length
* 82982 101845: gap of 100 bp
* 101846 101945: contig of 18864 bp in length
* 101946 122755: gap of 100 bp
* 122756 122855: contig of 20810 bp in length
* 122856 142457: gap of 100 bp
* 142458 142557: contig of 19602 bp in length
* 142558 167357: gap of 100 bp
* 167358 167357: contig of 24800 bp in length.

```

## FEATURES

```

source
1..167357
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chr="20"
/map="20"
/clone="RP11-561P16"
/clone_lib="RP11-561P16 Human Male BAC"
1..1216
/note="assembly_fragment"
1317..2985
/note="assembly_fragment"
3086..4628
/note="assembly_fragment"
4729..7507
/note="assembly_fragment"
7508..11292
/note="assembly_fragment"
11293..14213
/note="assembly_fragment"

```

```

misc_feature 14314..17828
/note="assembly_fragment"
misc_feature 17929..21068
/note="assembly_fragment"
misc_feature 21169..25035
/note="assembly_fragment"
misc_feature 25136..28331
/note="assembly_fragment"
misc_feature 28432..33595
/note="assembly_fragment"
misc_feature 33696..40346
/note="assembly_fragment"
misc_feature 40447..43807
/note="assembly_fragment"
misc_feature 43908..48727
/note="assembly_fragment"
misc_feature 48828..56083
/note="assembly_fragment"
misc_feature 56184..67055
/note="assembly_fragment"
misc_feature 67156..82881
/note="assembly_fragment"
misc_feature 82982..101845
/note="assembly_fragment"
misc_feature 101946..122755
/note="assembly_fragment"
misc_feature 122856..142457
/note="assembly_fragment"
misc_feature 142558..167357
/note="assembly_fragment"

ORIGIN
Query Match 5.5% Score 234; DB 2; Length 167357;
Best Local Similarity 100.0%; Pred.No. 2.4e-114; Indels 0; Gaps 0;
Matches 234; Conservative 0; Mismatches 0;

Qy 2951 CTGAAAATTTTCAGGCTACTGTACTGTAGTCTGTGGAGCAACAGGCTGTGGCT 3010
Db 113360 CTGAAAATTTTCAGGCTACTGTACTGTAGTCTGTGGAGCAACAGGCTGTGGCT 113301

Qy 3011 GGTGTACTGTATCCAGCAATCTGCGAAAGGGAAATGCATAGAGGTTCTATAAAGGAC 3070
Db 113300 GGTGTACTGTATCCAGCAATCTGCGAAAGGGAAATGCATAGAGGTTCTATAAAGGAC 113241

Qy 3071 CAGTGAAGATGCTTTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAATT 3130
Db 113240 CAGTGAAGATGCTTTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAATT 113181

Qy 3131 CCAGCATGTTCTAGAGGACAGCAGATACACTGGTCTTTTCTATCTACTGTCAG 3184
Db 113180 CCAGCATGTTCTAGAGGACAGCAGATACACTGGTCTTTTCTATCTACTGTCAG 113127

```

## RESULT 22

```

AP002898 206647 bp DNA linear PRI 24-OCT-2000
LOCUS Homo sapiens genomic DNA, chromosome 20p, clone:13N6, complete
DEFINITION sequence.
ACCESSION AP002898
VERSION AP002898.1 GI:10998163
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 206647)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens 206,647 genomic DNA of 20p

```

```
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
ORIGIN
Query Match
Best Local Similarity
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
/db_xref="taxon:9606"
5.5%; Score 234; DB 6; Length 207433;
Pred. No. 2.3e-114;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2951 CTGAAAATTGTTTCAGGCTACTCTACCTGTAGTCATTGTTGGAGCAACCCAGGCTGTGGCT 3010
9813 CTGAAAATTGTTTCAGGCTACTCTACCTGTAGTCATTGTTGGAGCAACCCAGGCTGTGGCT 9872
3011 GGTGTACTGATCCCGACCAATCTGCGCAAGCCCTACAGGAAATTCATCCACAGCCCTGCTCAATT 3130
9873 GGTGTACTGATCCCGACCAATCTGCGCAAGCCCTACAGGAAATTCATCCACAGCCCTGCTCAATT 9992
3131 CCAGCATGTGTCTAGAGGACGACGATACAACTGGTCTTTTCATTTCACGTGTCCAG 3184
9993 CCAGCATGTGTCTAGAGGACGACGATACAACTGGTCTTTTCATTTCACGTGTCCAG 10046
RESULT 24
H009384S11
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SEGMENT
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
Location/Qualifiers
1..3589
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="20"
/map="20p13"
1585..1805
/gene="ATR"
/number=12
exon
ORIGIN
Query Match
Best Local Similarity
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
5.2%; Score 223; DB 9; Length 3589;
Pred. No. 3.3e-108;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1870 AGCACCATGTATGTTTCGGTGGTTTCAATAGTCTTCCTCCAGGACATCCTGTATTC 1929
1583 AGCACCATGTATGTTTCGGTGGTTTCAATAGTCTTCCTCCAGGACATCCTGTATTC 1642
1930 ACCTCGGACAGTGTGATCGCATCGAGTACCGCCCTTGTTCAGGACGAGGACCTGGT 1989
Published Only in DataBase (2000)
2 (bases 1 to 206647)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (22-OCT-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Overlapping Clones: AL132773, AL109804
Cytogenetic Position: 20p
STS Markers: D20S828, D20S813, stSG408, A005005, stS-H22126,
D20S867, D20S752
Additional author information
Hayflick, S.J.
Oregon Health Sciences University, Molecular & Medical Genetics
Dept. 3131 SW Sam Jackson Park Road, MC-L103, BSA-4596, Portland OR
97201-3098, USA
phone:1-503-494-6866, fax:1-503-494-6886
e-mail: hayflick@ohsu.edu.
Location/Qualifiers
1..206647
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="20"
/map="20p"
/clones="13N6"
ORIGIN
Query Match
Best Local Similarity
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
5.5%; Score 234; DB 9; Length 206647;
Pred. No. 2.3e-114;
Conservative 0; Mismatches 0; Indels 0; Gaps 0;
2951 CTGAAAATTGTTTCAGGCTACTCTACCTGTAGTCATTGTTGGAGCAACCCAGGCTGTGGCT 3010
61933 CTGAAAATTGTTTCAGGCTACTCTACCTGTAGTCATTGTTGGAGCAACCCAGGCTGTGGCT 61992
3011 GGTGTACTGATCCCGACCAATCTGCGCAAGCCCTACAGGAAATTCATCCACAGCCCTGCTCAATT 3070
61993 GGTGTACTGATCCCGACCAATCTGCGCAAGCCCTACAGGAAATTCATCCACAGCCCTGCTCAATT 62052
3071 CAGTGAAGATGCTTCCAGAGCCCTACAGGAAATTCATCCACAGCCCTGCTCAATT 3130
62053 CAGTGAAGATGCTTCCAGAGCCCTACAGGAAATTCATCCACAGCCCTGCTCAATT 62112
3131 CCAGCATGTGTCTAGAGGACGACGATACAACTGGTCTTTTCATTTCACGTGTCCAG 3184
62113 CCAGCATGTGTCTAGAGGACGACGATACAACTGGTCTTTTCATTTCACGTGTCCAG 62166
RESULT 23
AX326809
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
Location/Qualifiers
1..207433
/organism="Homo sapiens"
/mol_type="unassigned DNA"
AX326809
Sequence 5 from Patent WO0178894.
AX326809
AX326809.1 GI:18097525
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Keith, T.
Novel human gene relating to respiratory diseases, obesity, and
inflammatory bowel disease
Patent: WO 0178894-A 5 25-OCT-2001;
Genome Therapeutics Corp. (US)
Location/Qualifiers
1..207433
/organism="Homo sapiens"
/mol_type="unassigned DNA"
```



```
misc_feature 69213..86049
/note="assembly_fragment"
misc_feature 86150..107477
/note="assembly_fragment"
misc_feature 107578..131094
/note="assembly_fragment"
misc_feature 131195..155869
/note="assembly_fragment"

ORIGIN
Query Match 5.2%; Score 223; DB 2; Length 155869;
Best Local Similarity 100.0%; Pred. No. 2.1e-108; Indels 0; Gaps 0;
Matches 223; Conservative 0; Mismatches 0;

Qy 1870 AGCACCATGATGTTGCGGTGTTCAATAGTCTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGAT 1929
Db 10396 AGCACCATGATGTTGCGGTGTTCAATAGTCTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGAT 10337

Qy 1930 ACCTCGGACAGTGTGATGCGCATCGGAGTGAACCGCTGTTTATAGCAGGACCTGGT 1989
Db 10336 ACCTCGGACAGTGTGATGCGCATCGGAGTGAACCGCTGTTTATAGCAGGACCTGGT 10277

Qy 1990 ATTCCGGTGTGTGGAAACACAGGCTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGAT 2049
Db 10276 ATTCCGGTGTGTGGAAACACAGGCTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGAT 10217

Qy 2050 GAACAGAGAAAGTTAAATCAGAATGTTTTCAAAAGAA 2092
Db 10216 GAACAGAGAAAGTTAAATCAGAATGTTTTCAAAAGAA 10174

RESULT 26
H009384514 H009384514 1728 bp DNA linear PRI 14-DEC-2000
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 16.
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 16.
ACCESSION AF218902
VERSION AF218902.1 GI:8118068
KEYWORDS
SEGMENT 14 of 27
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1728)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 1728)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
source
1..1728
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
926..1169
/gene="ATRN"
/number=8
exon
Query Match 5.2%; Score 221; DB 9; Length 1728;
Best Local Similarity 100.0%; Pred. No. 4.4e-107; Indels 0; Gaps 0;
Matches 221; Conservative 0; Mismatches 0;

Qy 1238 CAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTCATGGTGT 1297
Db 960 CAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTCATGGTGT 1019

Qy 1298 TGTGTACCCCTAAGGCAAGGAGCAGTATGTCAGTGGTTGGGCACCTCTGCACACATTTGA 1357
Db 1020 TGTGTACCCCTAAGGCAAGGAGCAGTATGTCAGTGGTTGGGCACCTCTGCACACATTTGA 1079

Qy 1358 CACTGAAGAATGCGCGAGTGGTCATGCTGGTCACTTTTGGTCACTGCCCTCTCTATGAT 1417
Db 1080 CACTGAAGAATGCGCGAGTGGTCATGCTGGTCACTTTTGGTCACTGCCCTCTCTATGAT 1139
```

```

QY      1418 ATATAAGCAATGTCAGGAATATGATTGG 1447
      |||
Db      1140 ATATAAGCAATGTCAGGAATATGATTGG 1169

RESULT 28
AC074061/c
LOCUS   Homo sapiens chromosome 2 clone RP11-770E8 map 2, LOW-PASS SEQUENCE
DEFINITION
ACCESSION AC074061
VERSION   AC074061.1 GI:9150468
KEYWORDS  HTG; HTGS PHASE0
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 63513)
          Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE     Homo sapiens chromosome 2, clone RP11-770E8
JOURNAL   Unpublished
AUTHORS   2 (bases 1 to 63513)
          Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
          Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
          Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
          Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
          Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
          Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
          Galagan,J., Gardyna,S., Glade,S., Goyette,M., Graham,L.,
          Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
          Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
          Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J.,
          Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
          McCarthy,M., McEwan,P., McGuck,A., McKernan,K., McPheeters,R.,
          Meldrim,J., Meneus,L., Mikova,T., Miranda,C., Mlenga,V., Morrow,J.,
          Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
          O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
          Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
          Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
          Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
          Tesfaye,S., Theodore,J., Tirrell,A., Traversan,M., Trigilio,J.,
          Vasiliev,H., Viel,R., Vo.A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
          Young,G., Zainoun,J., Zimmer,A. and Zody,M.
          Direct Submission
SUBMITTED (13-JUL-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WBIR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5682
Center clone name: 770_E_8
-----
* NOTE: This record contains 72 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
*
* 1      780: contig of 780 bp in length
* 781    880: gap of 100 bp
* 881    1658: contig of 778 bp in length
* 1659   1758: gap of 100 bp

```





```

RESULT 30
H009384S09
LOCUS          3038 bp      DNA      linear      PRI 14-DEC-2000
DEFINITION    Homo sapiens attractin precursor (ATRN) gene, exons 9 and 10.
ACCESSION     AF218897
VERSION       AF218897.1 GI:8118063
KEYWORDS
SEGMENT
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens

REFERENCE
AUTHORS      Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
TITLE        Secreted and membrane attractin result from alternative splicing of
              the human ATRN gene
JOURNAL       Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE      20283944
PUBMED       10811918
REFERENCE
AUTHORS      Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
              Schlossman, S.F. and Duke-Cohan, J.S.
TITLE        Direct Submission
JOURNAL       Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
              Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
              USA
FEATURES
source       Location/Qualifiers
              1..3038
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="20"
              /map="20p13"
              /number=9
              /gene="ATRN"
              /number=9
              2047..2201
              /gene="ATRN"
              /number=10
              exon
              1114..1297
              /number=9
              exon
              2047..2201
              /number=10
              ORIGIN
Query Match      4.3%; Score 186; DB 9; Length 3038;
Best Local Similarity 100.0%; Pred. No. 3.6e-88;
Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1447 GATAGACACATGGAGTATTATACACACCCAGGGTCCCTTGTGCAAGGGGTACGGC 1506
      |||
Db 1113 GATAGAACACATGGAGTATTATACACACCCAGGGTCCCTTGTGCAAGGGGTACGGC 1172

QY 1507 CATAGCAGTGTTCAGACCATAGGACCCAGGGCCCTATACGTTTCATGGTGGCTACAAGGCT 1566
      |||
Db 1173 CATAGCAGTGTTCAGACCATAGGACCCAGGGCCCTATACGTTTCATGGTGGCTACAAGGCT 1232

QY 1567 TTCAGTGCATATAGTACCGGCTTGCAGATATCTTACCGATATGATGTGGATACCCAG 1626
      |||
Db 1233 TTCAGTGCATATAGTACCGGCTTGCAGATATCTTACCGATATGATGTGGATACCCAG 1292

QY 1627 ATGTGG 1632
      |||
Db 1293 ATGTGG 1298

RESULT 31
H009384S06
LOCUS          1032 bp      DNA      linear      PRI 14-DEC-2000
DEFINITION    Homo sapiens attractin precursor (ATRN) gene, exon 6.
ACCESSION     AF218894
VERSION       AF218894.1 GI:8118060
KEYWORDS
SEGMENT
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens

REFERENCE
AUTHORS      Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
              Schlossman, S.F. and Duke-Cohan, J.S.
TITLE        Direct Submission
JOURNAL       Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
              Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
              USA
FEATURES
source       Location/Qualifiers
              1..1032
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="20"
              /map="20p13"
              /number=6
              /gene="ATRN"
              /number=6
              ORIGIN
Query Match      4.0%; Score 173; DB 9; Length 1032;
Best Local Similarity 100.0%; Pred. No. 4.5e-81;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 943 GGTCTGTGATGTTCAGTTCCTGTACAGCTAACAGTCAATTTTGGACTCGAGAGGAATAT 1002
      |||
Db 528 GGTCTGTGATGTTCAGTTCCTGTACAGCTAACAGTCAATTTTGGACTCGAGAGGAATAT 587

QY 1003 TCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATATATGCG 1062
      |||
Db 588 TCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATATATGCG 647

QY 1063 GTTGTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTTCTAGCGTA 1115
      |||
Db 648 GTTGTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTTCTAGCGTA 700

RESULT 32
H009384S05
LOCUS          1232 bp      DNA      linear      PRI 14-DEC-2000
DEFINITION    Homo sapiens attractin precursor (ATRN) gene, exon 5.
ACCESSION     AF218893
VERSION       AF218893.1 GI:8118059
KEYWORDS
SEGMENT
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens

REFERENCE
AUTHORS      Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
              and Duke-Cohan, J.S.
TITLE        Secreted and membrane attractin result from alternative splicing of
              the human ATRN gene
JOURNAL       Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE      20283944
PUBMED       10811918
REFERENCE
AUTHORS      Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
              Schlossman, S.F. and Duke-Cohan, J.S.
TITLE        Direct Submission
JOURNAL       Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
              Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
              USA

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
20283944
PUBMED 10811918
REFERENCE
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
source       Location/Qualifiers
              1..1032
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
              /chromosome="20"
              /map="20p13"
              /number=6
              /gene="ATRN"
              /number=6
              exon
              529..697
              /number=6
              ORIGIN
Query Match      4.0%; Score 173; DB 9; Length 1032;
Best Local Similarity 100.0%; Pred. No. 4.5e-81;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 943 GGTCTGTGATGTTCAGTTCCTGTACAGCTAACAGTCAATTTTGGACTCGAGAGGAATAT 1002
      |||
Db 528 GGTCTGTGATGTTCAGTTCCTGTACAGCTAACAGTCAATTTTGGACTCGAGAGGAATAT 587

QY 1003 TCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATATATGCG 1062
      |||
Db 588 TCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATATATGCG 647

QY 1063 GTTGTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTTCTAGCGTA 1115
      |||
Db 648 GTTGTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTTCTAGCGTA 700

RESULT 32
H009384S05
LOCUS          1232 bp      DNA      linear      PRI 14-DEC-2000
DEFINITION    Homo sapiens attractin precursor (ATRN) gene, exon 5.
ACCESSION     AF218893
VERSION       AF218893.1 GI:8118059
KEYWORDS
SEGMENT
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens

REFERENCE
AUTHORS      Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
              and Duke-Cohan, J.S.
TITLE        Secreted and membrane attractin result from alternative splicing of
              the human ATRN gene
JOURNAL       Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE      20283944
PUBMED       10811918
REFERENCE
AUTHORS      Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
              Schlossman, S.F. and Duke-Cohan, J.S.
TITLE        Direct Submission
JOURNAL       Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
              Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
              USA

```

[illegible]

Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

# TITLE JOURNAL

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 12, 2000 this sequence version replaced gi:7582593.  
All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L5007

Center clone name: 561\_P16

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 156504 bases at least Q40

Consensus quality: 161679 bases at least Q30

Consensus quality: 163964 bases at least Q20

Insert size: 170000; agarose-fp

Insert size: 165357; sum-of-ctgts

Quality coverage: 4.5 in Q20 bases; agarose-fp

Quality coverage: 4.6 in Q20 bases; sum-of-ctgts

----- NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 1216: contig of 1216 bp in length
1217 1316: gap of 100 bp
1317 2985: contig of 1669 bp in length
2986 3085: gap of 100 bp
3086 4628: contig of 1543 bp in length
4629 4728: gap of 100 bp
4729 7507: contig of 2779 bp in length
7508 7607: gap of 100 bp
7608 11292: contig of 3685 bp in length
11293 11392: gap of 100 bp
11393 14213: contig of 2821 bp in length
14214 14313: gap of 100 bp
14314 17828: contig of 3515 bp in length
17829 17928: gap of 100 bp
17929 21068: contig of 3140 bp in length
21069 21168: gap of 100 bp
21169 25035: contig of 3867 bp in length
25036 25135: gap of 100 bp
25136 28331: contig of 3196 bp in length
28332 28431: gap of 100 bp
28432 33595: contig of 5164 bp in length
33596 33696: gap of 100 bp
33697 40346: contig of 6651 bp in length
40347 40446: gap of 100 bp
40447 43807: contig of 3361 bp in length
43808 43907: gap of 100 bp
43908 48727: contig of 4820 bp in length
48728 48828: gap of 100 bp
48829 56183: contig of 7256 bp in length
56184 56184: gap of 100 bp
56185 67065: contig of 10882 bp in length
67066 67165: gap of 100 bp

```

```

* 67166 82881: contig of 15716 bp in length
* 82882 82981: gap of 100 bp
* 82982 101845: contig of 18864 bp in length
* 101846 101945: gap of 100 bp
* 101946 122755: contig of 20810 bp in length
* 122756 122855: gap of 100 bp
* 122856 142457: contig of 19602 bp in length
* 142458 142557: gap of 100 bp
* 142558 167357: contig of 24800 bp in length.

```

## FEATURES

### source

```

1..167357
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20"
/clone_lib="RPC1-11 Human Male BAC"

```

```

1..1216
/note="assembly_fragment"

```

```

misc_feature
1317..2985
/note="assembly_fragment"

```

```

misc_feature
3086..4628
/note="assembly_fragment"

```

```

misc_feature
4729..7507
/note="assembly_fragment"

```

```

misc_feature
7608..11292
/note="assembly_fragment"

```

```

misc_feature
11393..14213
/note="assembly_fragment"

```

```

misc_feature
14314..17828
/note="assembly_fragment"

```

```

misc_feature
17929..21068
/note="assembly_fragment"

```

```

misc_feature
21169..25035
/note="assembly_fragment"

```

```

misc_feature
25136..28331
/note="assembly_fragment"

```

```

misc_feature
28432..33595
/note="assembly_fragment"

```

```

misc_feature
33696..40346
/note="assembly_fragment"

```

```

misc_feature
40447..43807
/note="assembly_fragment"

```

```

misc_feature
43908..48727
/note="assembly_fragment"

```

```

misc_feature
48828..56083
/note="assembly_fragment"

```

```

misc_feature
56184..67065
/note="assembly_fragment"

```

```

misc_feature
67166..82881
/note="assembly_fragment"

```

```

misc_feature
82982..101845
/note="assembly_fragment"

```

```

misc_feature
101946..122755
/note="assembly_fragment"

```

```

misc_feature
122856..142457
/note="assembly_fragment"

```

```

misc_feature
142558..167357
/note="assembly_fragment"

```

## ORIGIN

```

Query Match          3.8%   Score 161;   DB 2;   Length 167357;
Best Local Similarity 100.0%;   Pred.No. 7.7e-75;
Matches 161;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY 3641 CAGCTGGAACCCAGGCTGGAGAAGAGATGCTGTGTTTCAAAAACCAACATTAGGAGT 3700
      |||||
DB 2352 CAGCTGGAACCCAGGCTGGAGAAGAGATGCTGTGTTTCAAAAACCAACATTAGGAGT 2411

```

QY 3701 ACAAGATAGTTTCTTAATGAGAGTTGATTTTCGACACCCCAATATCACTTTCT 3760  
 Db 2412 ACAAGATAGTTTCTTAATGAGAGTTGATTTTCGACACCCCAATATCACTTTCT 2471  
 QY 3761 TTGTTATGTCAGTAATTCACCTGGCCCAATCAAAATTCAG 3801  
 Db 2472 TTGTTATGTCAGTAATTCACCTGGCCCAATCAAAATTCAG 2512

RESULT 35  
 AC017113 206329 bp DNA linear HTG 07-JUL-2000  
 LOCUS Homo sapiens chromosome 20 clone RP11-574H7, WORKING DRAFT  
 DEFINITION SEQUENCE, 31 unordered pieces.  
 ACCESSION AC017113  
 VERSION AC017113.4 GI:8570339  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 206329)  
 Waterston,R.H.  
 The sequence of Homo sapiens clone  
 Unpublished  
 2 (bases 1 to 206329)  
 Waterston,R.H.  
 Direct Submission  
 Submitted (09-DEC-1999) Genome Sequencing Center, Washington  
 University School of Medicine, 444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 On Jun 17, 2000 this sequence version replaced gi:6855220.

----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: http://genome.wustl.edu/gsc/index.shtml

----- Project Information -----  
 Center project name: H.NH0574H07

----- Summary Statistics -----  
 Sequencing vector: M13; 91%  
 Sequencing vector: plasmid; 9%  
 Chemistry: Dye-terminator; Big Dye; % of reads  
 Chemistry: Dye-terminator; Big Dye; % of reads

Assembly program: Phrap; version 0.990319  
 Consensus quality: 184921 bases at least Q40  
 Consensus quality: 192397 bases at least Q30  
 Consensus quality: 195862 bases at least Q20

Insert size: 237000; agarose-fp  
 Insert size: 203329; sum-of-contigs

Quality coverage: 3.84 in Q20 bases; agarose-fp  
 Quality coverage: 3.66 in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 31 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1539: contig of 1539 bp in length  
 \* 1540 1639: gap of unknown length  
 \* 1640 3259: contig of 1820 bp in length  
 \* 3260 3359: gap of unknown length  
 \* 3360 5400: contig of 2041 bp in length  
 \* 5401 5500: gap of unknown length  
 \* 5501 7783: contig of 2283 bp in length  
 \* 7784 7883: gap of unknown length  
 \* 7884 11241: contig of 3358 bp in length  
 \* 11242 11341: gap of unknown length  
 \* 11342 13482: contig of 2141 bp in length

13483 13582: gap of unknown length  
 \* 13583 16915: contig of 3333 bp in length  
 \* 16916 17015: gap of unknown length  
 \* 17016 21346: contig of 4331 bp in length  
 \* 21347 21447: gap of unknown length  
 \* 21448 25103: contig of 3657 bp in length  
 \* 25104 25203: gap of unknown length  
 \* 25204 28415: contig of 3212 bp in length  
 \* 28416 28516: gap of unknown length  
 \* 28517 32873: contig of 4358 bp in length  
 \* 32874 32974: gap of unknown length  
 \* 32975 36496: contig of 3523 bp in length  
 \* 36497 36596: gap of unknown length  
 \* 36597 40649: contig of 4053 bp in length  
 \* 40650 40749: gap of unknown length  
 \* 40750 44878: contig of 4129 bp in length  
 \* 44879 44978: gap of unknown length  
 \* 44979 50353: contig of 5375 bp in length  
 \* 50354 50453: gap of unknown length  
 \* 50454 55946: contig of 5493 bp in length  
 \* 55947 56047: gap of unknown length  
 \* 56048 61421: contig of 5374 bp in length  
 \* 61422 67819: contig of 6299 bp in length  
 \* 67820 67919: gap of unknown length  
 \* 67920 72327: contig of 4408 bp in length  
 \* 72328 72428: gap of unknown length  
 \* 72429 77393: contig of 4866 bp in length  
 \* 77394 82626: contig of 5233 bp in length  
 \* 82627 82727: gap of unknown length  
 \* 82728 90186: contig of 7460 bp in length  
 \* 90187 90286: gap of unknown length  
 \* 90287 96584: contig of 6298 bp in length  
 \* 96585 96685: gap of unknown length  
 \* 96686 106019: contig of 9335 bp in length  
 \* 106020 113630: contig of 7411 bp in length  
 \* 113631 113630: gap of unknown length  
 \* 113631 125151: contig of 11521 bp in length  
 \* 125152 125251: gap of unknown length  
 \* 125252 137677: contig of 12426 bp in length  
 \* 137678 137777: gap of unknown length  
 \* 137778 151263: contig of 13486 bp in length  
 \* 151264 151363: gap of unknown length  
 \* 151364 165370: contig of 14007 bp in length  
 \* 165371 165470: gap of unknown length  
 \* 165471 184704: contig of 19234 bp in length  
 \* 184705 184805: gap of unknown length  
 \* 184806 206329: contig of 21525 bp in length.

FEATURES  
 Location/Qualifiers

1..206329  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /clone="RP11-574H7"

ORIGIN

Query Match 3.8%; Score 161; DB 2; Length 206329;  
 Best Local Similarity 100.0%; Pred. No. 7.5e-75;  
 Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3641 CAGCTGGAACCCAGGCTGGAGAGAGATGCTGTTGTTTCAAAAACCAATTAAGGAGT 3700  
 Db 173247 CAGCTGGAACCCAGGCTGGAGAGAGATGCTGTTGTTTCAAAAACCAATTAAGGAGT 173306  
 QY 3701 ACAAGATAGTTTCTTAATGAGAGATTTGATTTTCGACACCCCAATATCACTTTCT 3760  
 Db 173307 ACAAGATAGTTTCTTAATGAGAGATTTGATTTTCGACACCCCAATATCACTTTCT 173366  
 QY 3761 TTGTTATGTCAGTAATTCACCTGGCCCAATCAAAATTCAG 3801



```

REFERENCE
Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE
Direct Submission
JOURNAL
Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
Location/Qualifiers
1..1674
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
810..947
/gene="ATRN"
/number=19
source
exon
ORIGIN
Query Match 3.3%; Score 141; DB 9; Length 1674;
Best Local Similarity 100.0%; Pred. No. 8.8e-64;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3182 CAGTTCCTCCATCAACGGCCACAGTAATGATCATCATCATCATCTGTGAGAGTGTG 3241
Db 807 CAGTTCCTCCATCAACGGCCACAGTAATGATCATCATCATCTGTGAGAGTGTG 866
Qy 3242 AGAACCTTGACACAGGCAAGCAGCTCGAGACCTGCATATCTGGTTCACGGTGATCCCA 3301
Db 867 AGAACCTTGACACAGGCAAGCAGCTCGAGACCTGCATATCTGGTTCACGGTGATCCCA 926
Qy 3302 CCAATGGAGGGAATGTCAGC 3322
Db 927 CCAATGGAGGGAATGTCAGC 947
RESULT 39
AC013324 118540 bp DNA linear HTG 13-JUL-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-388K24 map 2, LOW-PASS
SEQUENCE SAMPLING.
AC013324
VERSION AC013324.3 GI:9123914
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Biren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 2, clone RP11-388K24
Unpublished
2 (bases 1 to 118540)
Biren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Bouckhalt, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lieu, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrim, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced gi:6289165.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

```

```

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3787
Center clone name: 388_K_24
-----

```

```

* NOTE: This record contains 125 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

```

```

1
856: contig of 856 bp in length
857
956: gap of 100 bp
957
1821: contig of 863 bp in length
1822
1821: gap of 100 bp
1922
2776: contig of 855 bp in length
2777
2876: gap of 100 bp
2877
3767: contig of 891 bp in length
3768
3867: gap of 100 bp
3868
4709: contig of 842 bp in length
4710
4809: gap of 100 bp
4810
5855: contig of 846 bp in length
5856
5755: gap of 100 bp
5756
6627: contig of 872 bp in length
6628
6727: gap of 100 bp
6728
7586: contig of 859 bp in length
7587
7686: gap of 100 bp
7687
8541: contig of 855 bp in length
8542
8641: gap of 100 bp
8642
9524: contig of 883 bp in length
9525
9624: gap of 100 bp
9625
10510
10609: contig of 885 bp in length
10510
11474: contig of 865 bp in length
11475
11574: gap of 100 bp
11575
12441: contig of 867 bp in length
12442
12541: gap of 100 bp
12542
13409: contig of 868 bp in length
13410
13509: gap of 100 bp
13510
14439: contig of 930 bp in length
14440
14539: gap of 100 bp
14540
15433: contig of 894 bp in length
15434
15533: gap of 100 bp
15534
16442: contig of 909 bp in length
16443
16542: gap of 100 bp
16543
17421: contig of 879 bp in length
17422
17521: gap of 100 bp
17522
18380: contig of 859 bp in length
18381
18480: gap of 100 bp
18481
19345: contig of 865 bp in length
19346
19445: gap of 100 bp
19446
20327: contig of 882 bp in length
20328
20427: gap of 100 bp
20428
21297: contig of 870 bp in length
21298
21397: gap of 100 bp
21398
22055: contig of 658 bp in length
22056
22155: gap of 100 bp
22156
23005: contig of 850 bp in length
23006
23105: gap of 100 bp
23106
24016: contig of 911 bp in length
24017
24116: gap of 100 bp
24117
24872: contig of 856 bp in length
24873
25072: gap of 100 bp
24973

```







```

repeat_region 11683..11848
/note="L1 repeat: matches 3736..3898 of consensus"
repeat_region 11849..12157
/note="AluSg repeat: matches 1..309 of consensus"
repeat_region 12158..12182
/note="L1 repeat: matches 3712..3736 of consensus"
repeat_region 12200..12396
/note="L1MC repeat: matches 1702..1912 of consensus"
repeat_region 12417..12586
/note="L1MC5 repeat: matches 7758..7913 of consensus"
repeat_region 12634..12914
/note="AluX repeat: matches 6..312 of consensus"
repeat_region 12958..13404
/note="L1MC repeat: matches 1267..1711 of consensus"
repeat_region 13486..13733
/note="AluSg1 repeat: matches 36..285 of consensus"
repeat_region 14313..14415
/note="MER86 repeat: matches 89..196 of consensus"
repeat_region 14493..15186
/note="L1 repeat: matches 4409..5090 of consensus"
repeat_region 15187..15487
/note="AluX repeat: matches 12..312 of consensus"
repeat_region 15488..15539
/note="L1 repeat: matches 4358..4409 of consensus"
misc_feature 15537..15942
/note="match: STS: Em:R15876"
misc_feature 15537..15748
/note="match: STS: Em:G15134"
gene complement(15767..17243)
/gene="dJ581P3.1"
/pseudo
complement(15767..17243)
/gene="dJ581P3.1"
60kD (SAP61) pseudogene)
match: CDNAs: Em:X81789 Em:U08815
match: ESTs: Em:A112866 Em:T56818 Em:AW470436 Em:AA300516
Em:AA703023 Em:AA760210 Em:AA054241 Em:AA443032 Em:R17811
Em:AW493824 Em:H67476 Em:A1020542 Em:AA343902 Em:H89818
Em:AA308239 Em:AA311786 Em:A1541046 Em:AA239392
Em:AA694306 Em:AA086662 Em:R15877 Em:A1226265 Em:P22705
Em:A1155866 Em:C06258 Em:AA086076 Em:AA081903 Em:A1019200
Em:AV138764 Em:AA06960 Em:A1202796 Em:R00790 Em:AA308500
Em:AA0303841 Em:AV290458 Em:AV166565 Em:AA074650
Em:AA135375 Em:AA355911 Em:AA730707 Em:AA068453 Em:T27366
Em:AA306883 Em:AL035731 Em:AA992803 Em:AL035732
Em:AA854771 Em:AA320502 Em:AA599259 Em:A1374964
Em:A1338280 Em:AA662262 Em:AA642724 Em:W83741 Em:AA662366
Em:AA463475 Em:W35845 Em:A1269312 Em:AL118208 Em:AA986711
Em:A1374716 Em:A1908435 Em:C82711 Em:AL039268 Em:AW246145
match: proteins: Tr:Q12874 Wp:FLH5 Tr:O59706 Tr:Q22469
Tr:O46106 Sw:P19736"
/pseudo
/codon_start=1
/evidence=not_experimental
/misc_feature 16008..16100
/note="match: GSS: Em:AQ279830"
/misc_feature 16104..16193
/note="match: GSS: Em:AQ074332"
/misc_feature complement(16881..16954)
/gene="dJ581P3.1"
/note="match: GSS: Em:AQ975997"
/pseudo
/misc_feature 16980..17146
/note="match: GSS: Em:AQ373489"
/misc_feature 17150..17265
/note="match: GSS: Em:AQ373489"
repeat_region 17389..17749
/note="L2 repeat: matches 1140..1536 of consensus"
repeat_region 17859..18030
/note="L2 repeat: matches 1791..1990 of consensus"
repeat_region 18405..18442
/note="19 copies 2 mer gt 97% conserved"

```

```

misc_feature complement(19050..19224)
/note="match: GSS: Em:AQ472547"
repeat_region 19051..19092
/note="21 copies 2 mer ag 90% conserved"
misc_feature complement(19093..19250)
/note="match: GSS: Em:AQ271994"
repeat_region 19142..19175
Query Match 3.1%; Score 131; DB 9; Length 73666;
Best Local Similarity 100.0%; Pred. No. 1.5e-58;
Matches 131; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 607 AGTGGCCTCATTTGTTCTTGAGAGATGCAATGAGATGTCCCTGAGGTGTGTTCCACA 666
|||||
Db 71998 AGTGGCCTCATTTGTTCTTGAGAGATGCAATGAGATGTCCCTGAGGTGTGTTCCACA 72057
|||||
QY 667 TAGGTTATGCTTGTGCTGCAATTTTATGTCGTGCTATTAATTTGACTGGATTTAAT 726
|||||
Db 72058 TAGGTTATGCTTGTGCTGCAATTTTATGTCGTGCTATTAATTTGACTGGATTTAAT 72117
|||||
QY 727 ATTACTTACAG 737
|||||
Db 72118 ATTACTTACAG 72128
|||||
RESULT 42
H009384S27 5790 bp DNA linear PRI 14-DEC-2000
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 30 and complete
DEFINITION cds, alternatively spliced.
ACCESSION AF218915
VERSION AF218915.1 GI:8118081
KEYWORDS 27 of 27
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 5790)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlozman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 5790)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlozman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
source
1..5790
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/map="20p13"
/order(AF218889..1:745..1576,AF218890..1:1..3711,
AF218891..1..3058,AF218892..1:1..1152,AF218893..1:1..1232,
AF218894..1:1..1032,AF218895..1:1..1575,AF218896..1:1..2018,
AF218897..1:1..3038,AF218898..1:1..1552,AF218899..1:1..3589,
AF218900..1:1..1267,AF218901..1..3112,AF218902..1:1..1728,
AF218903..1:1..3257,AF218904..1:1..1674,AF218905..1:1..1675,
AF218906..2:1..966,AF218907..1:1..1781,AF218908..1:1..1704,
AF218909..1:1..1976,AF218910..1:1..1364,AF218911..1:1..2214,
AF218912..1:1..1556,AF218913..1:1..1592,AF218914..1:1..1394,
1..5060)
/gene="ATRN"
join(AF218889..1:745..1233,AF218890..1:2877..2960,
AF218891..1:810..923,AF218892..1:818..946,

```

```
AF218893.1:621..826,AF218894.1:529..697,
AF218895.1:11343..1433,AF218896.1:1926..1169,
AF218897.1:11114..1297,AF218897.1:2047..2201,
AF218898.1:1659..743,AF218899.1:1585..1805,
AF218900.1:1728..849,AF218901.1:440..578,
AF218901.1:2163..2353,AF218902.1:533..752,
AF218903.1:1441..1626,AF218903.1:2189..2422,
AF218904.1:810..947,AF218905.1:833..926,
AF218906.2:49..101,AF218907.1:802..896,
AF218911.1:1348..951,AF218909.1:392..549,
AF218911.1:1348..1409,AF218912.1:748..826,
AF218913.1:1747..854,AF218914.1:735..849,660..5060)
/gene="ATRN"
/product="membrane attractin precursor"
join(AF218899.1:824..1233,AF218890.1:2877..2960,
AF218891.1:810..923,AF218892.1:818..946,
AF218893.1:621..826,AF218894.1:529..697,
AF218895.1:111343..1433,AF218896.1:1926..1169,
AF218897.1:11114..1297,AF218897.1:2047..2201,
AF218898.1:1659..743,AF218899.1:1585..1805,
AF218900.1:1728..849,AF218901.1:440..578,
AF218901.1:2163..2353,AF218902.1:533..752,
AF218903.1:1441..1626,AF218903.1:2189..2422,
AF218904.1:810..947,AF218905.1:833..926,
AF218906.2:49..101,AF218907.1:802..896,
AF218911.1:1348..951,AF218909.1:392..549,
AF218911.1:1348..1409,AF218912.1:748..826,
AF218913.1:1747..854,AF218914.1:735..849,660..784)
/gene="ATRN"
/note="Method: conceptual translation with partial peptide
sequencing"
/codon_start=1
/product="membrane attractin precursor"
/protein_id="AA72881.1"
/db_xref="GI:8118082"
/translation="MVAAATAEARLRRTAATAAAGRGSGPHWDVTRAGRGIG
AGLEPRLSPLEPRLLLGLLSPRLLLPCEAAASAAASGSAASAKED
RPGVNGCNFGTQCPCAGVGEQCQHCGRFRFGSSGFTDGFNGYKTKCTM
LIEQPNRINELNRPATECSDEHLYVDGDSIYAPLVAASGLIYPERDGNVTE
VWATGVALHFFSDAANTLGFNITVDFMPCNCSGRGECKISNDVCECGEN
WKGACDI PHTDNCNCPHRLGICNSDVRCSCFSDWQPGCVSPANCFVREY
SNLKLPRASHKAVNGNIMVVGVMFNHSDYNNLAYDLASREWLPLNRSVNVVR
YGHSLALYKOKIYMGKIDSTINELRVFHLHNSWVLLTPKAKEAVVGHSAH
IVTLKNGRVMLVIFGHCPLYGISNVDYDLDTQMTLSILHTQALVQGGYHSSVD
HRTALYVHGKYAFSANKYRLADLYRDYDVTQMTILKDRFFRYLHTAVTSGTM
LVFGNTHNDTNSHGAKCFSSDFMAYDIACDMSVLPDLHDLNRFHGSVALHNS
TMVYFGFNSLLIDILVFTSEOCDAHRSAAACIAGRCIRCVNWTGSSOCI SWALAT
DEQEKLSCEPFXRTLDHRCDDHTCYSTANTNDCHNCDCVPRNHSCEGQIS
IFRIENCFKPNMTYCNKTSKCRSCALDQNCWEPRNQECIALPEFCIGLWLVGNS
CLKITAKENYDNALFCRNHNALLSLTTQKRVFVLKOLRINQSSQSKLTLTPW
VGLRKINYSWCWDMSPFTNSLLQWPSSEPSDAGFCGLSEPSLTKAATCINFLN
GSVCERPAHSAKQCTPCALRTACGDTSGSSECMWNNKQCVDSNATVAFPPFQ
CMWYWTSTCPCPCSCVCTCHLQPGCGWCTDPSNTGKGCIEGSKYKGVKMPQ
APTGNFVPEPLNNSMCLDSRYNWSFIHCPACQNGHSCINQICEKCNLTGKH
CETCISGFYDPTNGGCKQCKCNCHASLNTWNTGKCFCTTKGVKDECOLCEVRY
QGNELRGCTYLLIDVFTSLCKDDRYITAINFVATDEQNRDLDMFINASKNPN
LNIITWAFSFGAQGEMBPVYKTNKEYDSFNEKFDPRHNPTTFVYVNSFTW
PQIKQIAFSQNSFMIDVQVFTFPFSCFLSLVAAYVWKIKOSCWASRRERQLREM
QOMASRPFASDEVDPEPDLIGGSIKTPKPIALEPFCGNKAAVLVSVFVLLPRG
LGGIPPPQOSGLAVASLVDISQMPVIYKESGAVNRKQPPAOPGTCTI"
join(AF218899.1:1073..1233,AF218890.1:2877..2960,
AF218891.1:810..923,AF218892.1:818..946,
AF218893.1:621..826,AF218894.1:529..697,
AF218895.1:11343..1433,AF218896.1:1926..1169,
AF218897.1:11114..1297,AF218897.1:2047..2201,
AF218898.1:1659..743,AF218899.1:1585..1805,
AF218900.1:1728..849,AF218901.1:440..578,
AF218901.1:2163..2353,AF218902.1:533..752,
AF218903.1:1441..1626,AF218903.1:2189..2422,
AF218904.1:810..947,AF218905.1:833..926,
AF218906.2:49..101,AF218907.1:802..896,
AF218911.1:1348..951,AF218909.1:392..549,
AF218911.1:1348..1409,AF218912.1:748..826,
AF218913.1:1747..854,AF218914.1:735..849,660..784)
```

mat\_peptide

```
AF218913.1:1747..854,AF218914.1:735..849,660..781)
/gene="ATRN"
/product="membrane attractin"
660..5060
/gene="ATRN"
/number=30
polyA_signal
5038..5043
/gene="ATRN"
ORIGIN
Query Match 3.0%; Score 127; DB 9; Length 5790;
Best Local Similarity 100.0%; Pred. No. 2.9e-56;
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4164 AGCTCTTGTCTGGCCAGCGCCCTGTGTGACATTTCTAGCAGATGCCATAGTGTACAA 4223
Db 658 AGGTCTTGTCTGGCCAGCGCCCTGTGTGACATTTCTAGCAGATGCCATAGTGTACAA 717
QY 4224 GGAGAAAGTCAGGAGCGCTGAGAAACCGGAAGCAGACAGCCCTTGCACACCTTGGGACCTG 4283
Db 718 GGAGAAAGTCAGGAGCGCTGAGAAACCGGAAGCAGACAGCCCTTGCACACCTTGGGACCTG 777
QY 4284 CATCTCA 4290
Db 778 CATCTCA 784
RESULT 43
AC017113/c
LOCUS Homo sapiens chromosome 20 clone RP11-574H7, WORKING DRAFT
DEFINITION SEQUENCE, 31 unordered pieces.
ACCESSION AC017113, 4 GI:8570339
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Fukuyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 206329)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 206329)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (09-DEC-1999) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Jun 17, 2000 this sequence version replaced gi:6855220.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0574H07
----- Summary Statistics -----
Sequencing vector: pUC19
Sequencing vector: plasmid; 9%
Chemistry: Dye-terminator Big Dye; 9% of reads
Chemistry: Dye-terminator Big Dye; 9% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 184921 bases at least Q40
Consensus quality: 192397 bases at least Q30
Consensus quality: 195862 bases at least Q20
Insert size: 237000; agarose-fp
Insert size: 203329; sum-of-contigs
Quality coverage: 3.84 in Q20 bases; agarose-fp
Quality coverage: 3.66 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
```

\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1539: contig of 1539 bp in length  
\* 1540 1639: gap of unknown length  
\* 3259: contig of 1620 bp in length  
\* 3359: gap of unknown length  
\* 3360 5400: contig of 2041 bp in length  
\* 5401 5500: gap of unknown length  
\* 5501 7783: contig of 2283 bp in length  
\* 7784 7883: gap of unknown length  
\* 7884 11241: contig of 3358 bp in length  
\* 11242 13482: contig of 2141 bp in length  
\* 13483 13582: gap of unknown length  
\* 13583 16915: contig of 3333 bp in length  
\* 16916 17015: gap of unknown length  
\* 17016 21346: contig of 4331 bp in length  
\* 21347 25103: contig of 3657 bp in length  
\* 25104 28415: contig of 3212 bp in length  
\* 28416 32873: contig of 4358 bp in length  
\* 32874 32973: gap of unknown length  
\* 32974 36496: contig of 3523 bp in length  
\* 36497 36596: gap of unknown length  
\* 36597 40749: contig of 4053 bp in length  
\* 40750 44878: contig of 4129 bp in length  
\* 44879 50353: contig of 5375 bp in length  
\* 50354 55946: contig of 5493 bp in length  
\* 55947 61420: contig of 5374 bp in length  
\* 61421 61520: gap of unknown length  
\* 61521 67819: contig of 6299 bp in length  
\* 67820 72327: contig of 4408 bp in length  
\* 72328 72427: gap of unknown length  
\* 72428 77293: contig of 4866 bp in length  
\* 77294 82626: contig of 5233 bp in length  
\* 82627 82726: gap of unknown length  
\* 82727 90186: contig of 7460 bp in length  
\* 90187 90286: gap of unknown length  
\* 90287 96584: contig of 6298 bp in length  
\* 96585 106019: contig of 9335 bp in length  
\* 106020 106119: gap of unknown length  
\* 106120 113330: contig of 7411 bp in length  
\* 113331 113630: gap of unknown length  
\* 113631 125151: contig of 11521 bp in length  
\* 125152 125251: gap of unknown length  
\* 125252 137677: contig of 12426 bp in length  
\* 137678 137777: gap of unknown length  
\* 137778 151263: contig of 13486 bp in length  
\* 151264 151363: gap of unknown length  
\* 151364 165370: contig of 14007 bp in length  
\* 165371 184704: gap of unknown length  
\* 184705 184804: gap of unknown length  
\* 184805 206329: contig of 21525 bp in length.

## FEATURES

source

1..206329  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="20"

## ORIGIN

Query Match 3.0%; Score 127; DB 2; Length 206329;  
Best Local Similarity 100.0%; Pred. No. 1.9e-56;  
Matches 127; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4164 AGGTCTTGTGTGGCAGCGCCCTGTGGACATTTCTCAGCAGATGCGATAGTGTACAA 4223  
Db 159328 AGGTCTTGTGTGGCAGCGCCCTGTGGACATTTCTCAGCAGATGCGATAGTGTACAA 159269

QY 4224 GGAGAGTCAGGAGCGGTGAGAAACCGGAAGCAGCAGCCCTGCACAGCTGGGACCTG 4283  
Db 159268 GGAGAGTCAGGAGCGGTGAGAAACCGGAAGCAGCAGCCCTGCACAGCTGGGACCTG 159209

QY 4284 CATCTGA 4290

Db 159208 CATCTGA 159202

## RESULT 44

H009384S26

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SEGMENT

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

exon

ORIGIN

Query Match

Best Local Similarity

Matches 119;

Conservative

Mismatches

Indels

Gaps

QY

Db

QY

Db

QY

Db

RESULT 45

```
H009384S03
LOCUS H009384S03 3058 bp DNA linear PRI 14-DEC-2000
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 3.
ACCESSION AF218891
VERSION AF218891.1 GI:8118057
KEYWORDS
SEGMENT
SOURCE 3 of 27
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3058)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 3058)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
source Location/Qualifiers
1..3058
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
810..923
/feature="ATRN"
/number=3
exon
Query Match 2.7%; Score 114; DB 9; Length 3058;
Best Local Similarity 100.0%; Pred. No. 3.4e-49;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 495 GCCAAATAGATATAGACTTCGTTCAATCATTTGCTACAGAGTGTAGTGGACCA 554
Db 810 GCCAAATAGATATAGACTTCGTTCAATCATTTGCTACAGAGTGTAGTGGACCA 869
Qy 555 TTTATATGTTATGATGGGACTCAATTATGACCGCTAGTTCGCTGATTAG 608
Db 870 TTTATATGTTATGATGGGACTCAATTATGACCGCTAGTTCGCTGATTAG 923
RESULT 46
AF273610 237 bp mRNA linear MAM 09-JUL-2000
LOCUS AF273610
DEFINITION Sus scrofa mahogany mRNA, partial cds.
ACCESSION AF273610
VERSION AF273610.1 GI:8980838
KEYWORDS
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 237)
AUTHORS Matteri, R.L. and Dyer, C.J.
TITLE Sus scrofa mahogany mRNA
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 237)
AUTHORS Matteri, R.L. and Dyer, C.J.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2000) Animal Physiology Research Unit, USDA -
Agricultural Research Service, University of Missouri, 920 East
Campus Dr., Room S-143 ASRC, Columbia, MO 65211, USA
FEATURES
Location/Qualifiers
```

```
1..237
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="brain: hypothalamus"
<1..>237
/function="possible role in appetite control and energy
regulation"
/codon_start=2
/product="transmembrane protein"
/protein_id="AA82297.1"
/db_xref="GI:8980838"
/translation="TDCYSCTANTNDCHWNCNHCVPNMHSCSTEGQISIAKYENCPRDN
PMYCNKKTSCRSQALDQNCQWPRNQECIALPE"
ORIGIN
Query Match 2.6%; Score 113; DB 4; Length 237;
Best Local Similarity 100.0%; Pred. No. 1.6e-48;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2245 AAGGATAACCCCATGTACTGTAAACAAGAACAGCAGCTGCAGGAGCTGTGCCCTGGAC 2304
Db 125 AAGGATAACCCCATGTACTGTAAACAAGAACAGCAGCTGCAGGAGCTGTGCCCTGGAC 184
Qy 2305 CAGAACTGCAGTGGGAGCCCGGATCAGGAGTGCATTGCCCTGCCGAAA 2357
Db 185 CAGAACTGCAGTGGGAGCCCGGATCAGGAGTGCATTGCCCTGCCGAAA 237
RESULT 47
H009384S25 1592 bp DNA linear PRI 14-DEC-2000
LOCUS H009384S25
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 28.
ACCESSION AF218913
VERSION AF218913.1 GI:8118079
KEYWORDS
SEGMENT 25 of 27
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1592)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 1592)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
FEATURES
source Location/Qualifiers
1..1592
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
747..854
/feature="ATRN"
/number=28
exon
Query Match 2.6%; Score 110; DB 9; Length 1592;
Best Local Similarity 100.0%; Pred. No. 5.4e-47;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 3941 AGCAACTTCTTCGAGATGCAACAGATGGCCAGCGTCCCTTGGCTCTGTAATGTCG 4000
Db 745 AGCAACTTCTTCGAGATGCAACAGATGGCCAGCGTCCCTTGGCTCTGTAATGTCG 804
QY 4001 CTTTGAACAGATGAGAGCTCTGATCTATTATGGGGAGTATAAG 4050
Db 805 CTTTGAACAGATGAGAGCTCTGATCTATTATGGGGAGTATAAG 854

RESULT 48
H009384S07
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 20.
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 20.
ACCESSION AF218905
VERSION AF218905.1 GI:8118071
KEYWORDS
SEGMENT
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1675)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 1675)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
FEATURES
Source
1..1675
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
333..926
/gene="ATRN"
/number=20
exon
1173 TAGATATGTCATCTTTGGCATATACAGG 1204
1403 TAGATATGTCATCTTTGGCATATACAGG 1434
ORIGIN
Query Match 2.2%; Score 94; DB 9; Length 1675;
Best Local Similarity 100.0%; Pred. No. 2.4e-38;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3323 CATGCAAGTGCATCGGACCGCTCTGTGCAACACCAACGCGGCAAGTCTTGCA 3382
Db 833 CATGCAAGTGCATCGGACCGCTCTGTGCAACACCAACGCGGCAAGTCTTGCA 892
QY 3383 CCACCAAGGGCGTCAAGGGGACGAGTCCAGCT 3416
Db 893 CCACCAAGGGCGTCAAGGGGACGAGTCCAGCT 926

RESULT 49
H009384S07
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 7.
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 7.
ACCESSION AF218895
VERSION AF218895.1 GI:8118061
KEYWORDS
SEGMENT
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1575)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 3711)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
FEATURES
Source
1..3711
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
QY 1113 GTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGTCTCTGTAACAATGGTTGT 1172
Db 1343 GTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGTCTCTGTAACAATGGTTGT 1402
QY 1173 TAGATATGTCATCTTTGGCATATACAGG 1204
Db 1403 TAGATATGTCATCTTTGGCATATACAGG 1434
ORIGIN
Query Match 2.1%; Score 92; DB 9; Length 1575;
Best Local Similarity 100.0%; Pred. No. 3e-37;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1113 GTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGTCTCTGTAACAATGGTTGT 1172
Db 1343 GTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGTCTCTGTAACAATGGTTGT 1402
QY 1173 TAGATATGTCATCTTTGGCATATACAGG 1204
Db 1403 TAGATATGTCATCTTTGGCATATACAGG 1434
ORIGIN
Query Match 2.2%; Score 94; DB 9; Length 1675;
Best Local Similarity 100.0%; Pred. No. 2.4e-38;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3323 CATGCAAGTGCATCGGACCGCTCTGTGCAACACCAACGCGGCAAGTCTTGCA 3382
Db 833 CATGCAAGTGCATCGGACCGCTCTGTGCAACACCAACGCGGCAAGTCTTGCA 892
QY 3383 CCACCAAGGGCGTCAAGGGGACGAGTCCAGCT 3416
Db 893 CCACCAAGGGCGTCAAGGGGACGAGTCCAGCT 926

RESULT 49
H009384S07
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 7.
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 7.
ACCESSION AF218895
VERSION AF218895.1 GI:8118061
KEYWORDS
SEGMENT
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1575)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 3711)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
FEATURES
Source
1..3711
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"

```

```

REFERENCE 1 (bases 1 to 1575)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 1575)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
FEATURES
Source
1..1575
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
1343..1433
/gene="ATRN"
/number=7
exon
1173 TAGATATGTCATCTTTGGCATATACAGG 1204
1403 TAGATATGTCATCTTTGGCATATACAGG 1434
ORIGIN
Query Match 2.1%; Score 92; DB 9; Length 1575;
Best Local Similarity 100.0%; Pred. No. 3e-37;
Matches 92; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1113 GTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGTCTCTGTAACAATGGTTGT 1172
Db 1343 GTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGTCTCTGTAACAATGGTTGT 1402
QY 1173 TAGATATGTCATCTTTGGCATATACAGG 1204
Db 1403 TAGATATGTCATCTTTGGCATATACAGG 1434
ORIGIN
Query Match 2.2%; Score 94; DB 9; Length 1675;
Best Local Similarity 100.0%; Pred. No. 2.4e-38;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3323 CATGCAAGTGCATCGGACCGCTCTGTGCAACACCAACGCGGCAAGTCTTGCA 3382
Db 833 CATGCAAGTGCATCGGACCGCTCTGTGCAACACCAACGCGGCAAGTCTTGCA 892
QY 3383 CCACCAAGGGCGTCAAGGGGACGAGTCCAGCT 3416
Db 893 CCACCAAGGGCGTCAAGGGGACGAGTCCAGCT 926

RESULT 50
H009384S02
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 2.
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 2.
ACCESSION AF218890
VERSION AF218890.1 GI:8118056
KEYWORDS
SEGMENT
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3711)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 3711)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
FEATURES
Source
1..3711
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"

```



LOCUS	H009384S24	1556 bp	DNA	linear	PRI 14-DEC-2000
DEFINITION	Homo sapiens attractin precursor (ATRN) gene, exon 27.				
ACCESSION	AF218912				
VERSION	AF218912.1 GI:8118078				
KEYWORDS	24 of 27				
SEGMENT	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1556)				
AUTHORS	Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.				
TITLE	Secreted and membrane attractin result from alternative splicing of the human ATRN gene				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)				
MEDLINE	20283944				
PUBMED	10811918				
REFERENCE	2 (bases 1 to 1556)				
AUTHORS	Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA				
FEATURES	Location/Qualifiers				
source	1..1556				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/chromosome="20"				
	/map="20p13"				
	748..826				
	/gene="ATRN"				
	/number=27				
exon					
ORIGIN					
Query Match	1.9%; Score 82; DB 9; Length 1556;				
Best Local Similarity	100.0%; Pred. No. 7.7e-32;				
Matches	82; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	3861 CAGTGTGTTCTCTCTTGGCTCCCTGGTGGCTGCTGTTGGAGATCAACAAAGTTG 3920				
Db	745 CAGTGTGTTCTCTCTTGGCTCCCTGGTGGCTGCTGTTGGAGATCAACAAAGTTG 804				
QY	3921 TTGGGCGCTCCAGACGTAGAGAG 3942				
Db	805 TTGGGCGCTCCAGACGTAGAGAG 826				
RESULT 53	H009384S20 1704 bp DNA linear PRI 14-DEC-2000				
LOCUS	Homo sapiens attractin precursor (ATRN) gene, exon 23.				
DEFINITION					
ACCESSION	AF218908				
VERSION	AF218908.1 GI:8118074				
KEYWORDS	20 of 27				
SEGMENT	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1704)				
AUTHORS	Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.				
TITLE	Secreted and membrane attractin result from alternative splicing of the human ATRN gene				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)				
MEDLINE	20283944				
PUBMED	10811918				
REFERENCE	2 (bases 1 to 1704)				
AUTHORS	Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA				
FEATURES	Location/Qualifiers				
source	1..1552				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/chromosome="20"				
	/map="20p13"				
	659..1552				
	/gene="ATRN"				
	/number=11				
exon					
ORIGIN					
Query Match	2.0%; Score 86; DB 9; Length 1552;				
Best Local Similarity	100.0%; Pred. No. 5.3e-34;				
Matches	86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1786 GCCTGTGACCGGTGTCAGTCTTCCAGACCTGATCTCCACCATGATGTCACAGATTT 1845				
Db	658 GCCTGTGACCGGTGTCAGTCTTCCAGACCTGATCTCCACCATGATGTCACAGATTT 717				
QY	1846 GGCCATTTCAGCAGTCTTACACACAG 1871				
Db	718 GGCCATTTCAGCAGTCTTACACACAG 743				
RESULT 52	H009384S24				
LOCUS	1552 bp DNA linear PRI 14-DEC-2000				
DEFINITION	Homo sapiens attractin precursor (ATRN) gene, exon 11.				
ACCESSION	AF218898				
VERSION	AF218898.1 GI:8118064				
KEYWORDS	10 of 27				
SEGMENT	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1552)				
AUTHORS	Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.				
TITLE	Secreted and membrane attractin result from alternative splicing of the human ATRN gene				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)				
MEDLINE	20283944				
PUBMED	10811918				
REFERENCE	2 (bases 1 to 1552)				
AUTHORS	Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA				
FEATURES	Location/Qualifiers				
source	1..1552				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/chromosome="20"				
	/map="20p13"				
	659..1552				
	/gene="ATRN"				
	/number=11				
exon					
ORIGIN					
Query Match	2.0%; Score 86; DB 9; Length 1552;				
Best Local Similarity	100.0%; Pred. No. 5.3e-34;				

TITLE Direct Submission  
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,  
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,  
USA

FEATURES  
source  
1..1704 Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="20p13"  
873..951  
/gene="ATRN"  
/number=23

exon

ORIGIN  
Query Match 1.8%; Score 79; DB 9; Length 1704;  
Best Local Similarity 100.0%; Pred.No. 3.2e-30;  
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3565 CAAACAGGATTGGACATGTTTCATCATGCTCCAGAAATTTCAACCTCAACATCACC 3624  
|||||  
DB 873 CAAACAGGATTGGACATGTTTCATCATGCTCCAGAAATTTCAACCTCAACATCACC 932  
|||||

QY 3625 TGGGCTGCAGTTTCTCAG 3643  
|||||  
DB 933 TGGGCTGCAGTTTCTCAG 951  
|||||

RESULT 54  
H009384S12  
LOCUS 1267 bp DNA linear PRI 14-DEC-2000  
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 13.  
ACCESSION AF218900  
VERSION AF218900.1 GI:8118066

KEYWORDS  
SEGMENT  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1267)  
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.  
and Duke-Cohan, J.S.  
TITLE Secreted and membrane attractin result from alternative splicing of  
the human ATRN gene  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)  
MEDLINE 20283944  
PUBMED 10811918

REFERENCE 2 (bases 1 to 1267)  
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,  
Schlossman, S.F. and Duke-Cohan, J.S.  
TITLE Direct Submission  
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,  
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,  
USA

FEATURES  
source  
1..1267 Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="20p13"  
728..849  
/gene="ATRN"  
/number=13

exon

ORIGIN  
Query Match 1.7%; Score 75; DB 9; Length 1267;  
Best Local Similarity 100.0%; Pred.No. 4.9e-28;  
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2140 ACAGCCACCAATGACTGCGCAATGTCGCAATGACCATTTGTCTCCCGAGCAACACAGC 2199  
|||||

Db 775 ACAGCCACCAATGACTGCGCAATGTCGCAATGACCATTTGTCTCCCGAGCAACACAGC 834  
|||||

QY 2200 TGCTCAGAAGGCCAG 2214  
|||||

Db 835 TGCTCAGAAGGCCAG 849  
|||||

RESULT 55  
H009384S23  
LOCUS 2214 bp DNA linear PRI 14-DEC-2000  
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 26.  
ACCESSION AF218911  
VERSION AF218911.1 GI:8118077

KEYWORDS  
SEGMENT  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2214)  
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.  
and Duke-Cohan, J.S.  
TITLE Secreted and membrane attractin result from alternative splicing of  
the human ATRN gene  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)  
MEDLINE 20283944  
PUBMED 10811918

REFERENCE 2 (bases 1 to 2214)  
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,  
Schlossman, S.F. and Duke-Cohan, J.S.  
TITLE Direct Submission  
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,  
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,  
USA

FEATURES  
source  
1..2214 Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="20p13"  
1348..1409  
/gene="ATRN"  
/number=26

exon

ORIGIN  
Query Match 1.5%; Score 65; DB 9; Length 2214;  
Best Local Similarity 100.0%; Pred.No. 1.2e-22;  
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3799 CAGATTGCTTCTCTCAGCAGCAGCAATTTATGGACCTGGTACAGTTCTTCTGTGACTTTC 3858  
|||||

Db 1345 CAGATTGCTTCTCTCAGCAGCAGCAATTTATGGACCTGGTACAGTTCTTCTGTGACTTTC 1404  
|||||

QY 3859 TTCAG 3863  
|||||

Db 1405 TTCAG 1409  
|||||

RESULT 56  
AB062913  
LOCUS 4284 bp mRNA linear ROD 16-JAN-2002  
DEFINITION Mesocricetus auratus Atrn mRNA for attractin, complete cds.  
ACCESSION AB062913  
VERSION AB062913.1 GI:16930100

KEYWORDS  
SOURCE Mesocricetus auratus (golden hamster)  
ORGANISM Mesocricetus auratus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
Mesocricetus.

REFERENCE 1

**AUTHORS** Kuramoto,T., Nomoto,T., Fujiwara,A., Mizutani,M., Sugimura,T. and Ushijima,T.  
**TITLE** Insertional mutation of the Attractin gene in the black tremor hamster  
**JOURNAL** Mamm Genome 13 (1), 36-40 (2002)  
**MEDLINE** 21635551  
**PUBLISHED** 11773967  
**REFERENCE** 2 (bases 1 to 4284)  
**AUTHORS** Kuramoto,T. and Ushijima,T.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center Research Institute, Carcinogenesis Division; Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@ncc.go.jp, URL:http://www.ncc.go.jp/research/rat-genome/, Tel:81-3-3542-2511, Fax:81-3-5565-1753)  
**FEATURES** Location/Qualifiers  
 source  
 1. .4284  
 /organism="Vesocricetus auratus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10036"  
 1. .4284  
 /gene="Atrn"  
 1. .4284  
 /gene="Atrn"  
 /codon\_start=1  
 /product="attractin"  
 /protein\_id="BAB72012.1"  
 /db\_xref="GI:16930101"  
 /translations="MVAAATAEALLRGSTAAAPAGRKVEQYRLCAATGAWRPGPR  
 AGLCPLWVSRAALPPPLLLLLSLLPLPREAABAASVAVSGSAAAEAKCEDRP  
 CVNGRCNPGTCGCVAGVGCOCCHGGRFRLTGGSGFVTDGPNKYKTKTWLI  
 EGQPNRLRPNRNFHFAECSDHLYVVDGDSIVAPLIAAFGLIVPERDGNETAPEVT  
 VTSGYALLHFFSDAAVNLGTGNTYFDMCPNCSGRGCKSSNSNTVECESENK  
 GESCDPHCTDNCGFPHRGICNASDVRGDCSPHWQPGCSIIPVPAQNSFWKREEYSD  
 HSLALHKDKIYMGCKIDSTGNVFNELRVFHHNESWLLTPKAKEQYAVVGSAAHIV  
 TLAALVHVMVLIYFGHCPLYGIVSIVQYEDLEKNTWISILQOGALVGGYGHGSVYDHR  
 RTKALYVHGKYKAFSANKYRLADDLRYVDVDTQMTILKDSRFRYHLTAIVSGTMLV  
 EGTNNDTSMHGAKCFSSDFMAYDIACDRNSVLRPELHHDVNRHSAVLNYS  
 YVFGGNSLLSVDLVFTSEQDAHRSEACVAGGIRCLMDTQSGRCSWELATE  
 QAEKLKSECFSKRTLDRCDQHTDCYCTANTNDCHNCHVFNHSCTEGGQISIS  
 KYDNCPCPNMYCNKKTSCRCALDQNCQWEPNQCIALPENICGNHVLGNSCL  
 KITTAKENYDNALSCRHNHFLASLTQKVELVLRKQLRMOSSQMSKLTITPWG  
 LRKINYSYCWEDMSPTNLSLQWMPSEPSDAGFCGILSEPTSGRLKAATCINPLNG  
 VCRPNHSAKQCTPCALRTAGCTSSSECCWCSNMKQCVDSNAYVASFPFGQCM  
 EWTMSCCPFGSGYCTCSHCLFQPGCGCTDPSNTGKCKIEGSKYKGIPIBPSQAS  
 TGNVYFQPLNNSMCLSDRYNWSFHCPCACQNGHSCINQISICKENLTGKHCE  
 TGISGFYDPTNGKQCPCKNGHSLCNTNTGKCTCTTGKVGDECCQCEVNRVQG  
 NPLKTCYTYLLIDYQTFSLQSDDRYTAINFATPDQNRDLMDFINASKNFNL  
 ITWATSPAGTQTEGVVSVKTIKEYKDSFNEKDFRNHNTIFFVVSFTWPI  
 KIOIARSOHNSMDLVQFVTFPFCFLSLLVAVVMKIKOSCAWRREQLIREMOQ  
 MASRPASVNALETDEEPDLIGGSIKTVPKPILALEPCFGKAAVLVSFVRLPRGL  
 GIPFGQSGLAVALVDISQMPVIVYKESGAVNRKQKQPPAQPGTCTI"  
**ORIGIN**  
 Query Match 1.5%; Score 64; DB 10; Length 4284;  
 Best Local Similarity 100.0%; Pred. No. 3.8e-22;  
 Matches 64; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1607 GATATGATGGATACCCAGATGGACCATCTTAAAGGACGCCGATTTTCGGTACT 1666  
 Db 1601 GATATGATGGATACCCAGATGGACCATCTTAAAGGACGCCGATTTTCGGTACT 1660  
 Qy 1667 TGCA 1670  
 Db 1661 TGCA 1664  
 RESULT 57  
 AF119821  
 LOCUS 4313 bp mRNA linear ROD 16-APR-1999  
 DEFINITION Mus musculus attractin (Mgca) mRNA; complete cds.  
 ACCESSION AF119821

**AF119821.1** GI:4585306  
**KEYWORDS** Mus musculus (house mouse)  
**SOURCE** Mus musculus  
**ORGANISM** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
**REFERENCE** 1 (bases 1 to 4313)  
**AUTHORS** Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A., Schlossman,S.P., Duke-Cohan,J.S. and Barsh,G.S.  
**TITLE** The mouse mahogany locus encodes a transmembrane form of human attractin  
**JOURNAL** Nature 398 (6723), 152-156 (1999)  
**MEDLINE** 99844160  
**PUBLISHED** 10086356  
**REFERENCE** 2 (bases 1 to 4313)  
**AUTHORS** Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A., Schlossman,S.P., Duke-Cohan,J.S. and Barsh,G.S.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (12-JAN-1999) HMI, Stanford, Beckman Center B271A, Stanford, CA 94305-5323, USA  
**FEATURES** Location/Qualifiers  
 source  
 1. .4313  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /chromosome="2"  
 1. .4313  
 /gene="Mgca"  
 /note="mahogany candidate"  
 27. .4313  
 /gene="Mgca"  
 /note="mahogany"  
 /codon\_start=1  
 /product="attractin"  
 /protein\_id="AAD25372.1"  
 /db\_xref="GI:4585307"  
 /translations="MVAAVAASAEARLRGTTATAAPAGRKGRHRCPTATGAWRPG  
 PRARCLPRVLRALPPPLPLPLFLSLLPLPREAEEAAVAAVAVSGSAAAEAKCEDR  
 PCVNGRCNPGTCGCVAGVGCOCCHGGRFRLTGGSGFVTDGPNKYKTKTWLI  
 LEGYPNAVLRPNRNFHFAECSDHLYVVDGDSIVAPLIAAFGLIVPERDGNETAPEV  
 TVTSYALHFFSDAAVNLGTGNTYFDMCPNCSGRGCKSSNSNTVECESENW  
 KGESCDPHCTDNCGFPHRGICNASDVRGDCSPHWQPGCSIIPVPAQNSFWKREEYS  
 DLKPLASHKDKIYMGCKIDSTGNVFNELRVFHHNESWLLTPKAKEQYAVVGSAAHIV  
 GHSIALHKDKIYMGCKIDSTGNVFNELRVFHHNESWLLTPKAKEQYAVVGSAAHIV  
 VTLASGRVVMVLIYFGHCPLYGIVSIVQYEDLEKNTWISILQOGALVGGYGHGSAYDD  
 RTKALYVHGKYKAFSANKYRLADDLRYVDVDTQMTILKDSRFRYHLTAIVSGTML  
 VFGTNDTSMHGAKCFSSDFMAYDIACDRNSVLRPELHHDVNRHSAVLNYS  
 MYVFGGNSLLSVDLVFTSEQDAHRSEACVAGGIRCLMDTQSGRCSWELATE  
 EQAEKLKSECFSKRTLDRCDQHTDCYCTANTNDCHNCHVFNHSCTEGGQISIS  
 AKYESCPDNPMYCNKKTSCRCALDQNCQWEPNQCIALPENICGNHVLGNSC  
 LKITTAKENYDNALSCRHNHFLASLTQKVELVLRKQLRMOSSQMSKLTITPWG  
 GLRKINYSYCWEDMSPTNLSLQWMPSEPSDAGFCGILSEPTSGRLKAATCINPLNG  
 SVCEPRNHSKQCTPCALRTAGCTSSSECCWCSNMKQCVDSNAYVASFPFGQC  
 MEWTMSCCPFGSGYCTCSHCLFQPGCGCTDPSNTGKCKIEGSKYKGIPIBPSQAS  
 SAGNVYFQPLNNSMCLSDRYNWSFHCPCACQNGHSCINQISICKENLTGKHCE  
 ETCISGFYDPTNGKQCPCKNGHSLCNTNTGKCTCTTGKVGDECCQCEVNRVYQ  
 GNPLKTCYTYLLIDYQTFSLQSDDRYTAINFATPDQNRDLMDFINASKNFNL  
 NITWATSPAGTQTEGVVSVKTIKEYKDSFNEKDFRNHNTIFFVVSFTWPI  
 KIOIARSOHNSMDLVQFVTFPFCFLSLLVAVVMKIKOSCAWRREQLIREMOQ  
 QMASRPASVNALETDEEPDLIGGSIKTVPKPILALEPCFGKAAVLVSFVRLPRGL  
 GIPFGQSGLAVALVDISQMPVIVYKESGAVNRKQKQPPAQPGTCTI"  
**ORIGIN**  
 Query Match 1.4%; Score 62; DB 10; Length 4313;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-21;  
 Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3130 TCACGATGTGTCTAGAGACGACGACATGACTGGTCTTTTCATTCTCCAGCTTGC 3189  
 Db 3153 TCACGATGTGTCTAGAGACGACGACATGACTGGTCTTTTCATTCTCCAGCTTGC 3212  
 Qy 3190 CA 3191

Db	3213	CA	3214	
RESULT 58	AB038388			
LOCUS	Rattus norvegicus Atrn mRNA for attractin, complete cds.			
DEFINITION	AB038388			
ACCESSION	AB038388.1	GI:12275311		
VERSION	attractin.			
KEYWORDS	Rattus norvegicus (Norway rat)			
SOURCE	Rattus norvegicus			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
REFERENCE	1 (sites)			
AUTHORS	Kuramoto, T., Kitada, K., Inui, T., Sasaki, Y., Ito, K., Hase, T., Kawaguchi, S., Ogawa, Y., Nakao, K., Barsh, G.S., Nagao, M., Ushijima, T. and Serikawa, T.			
TITLE	Attractin/mahogany/zitter plays a critical role in myelination of the central nervous system			
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)			
MEDLINE	21143347			
PUBMED	11209055			
REFERENCE	2 (bases 1 to 4500)			
AUTHORS	Kuramoto, T., Serikawa, T. and Ushijima, T.			
TITLE	Direct Submission			
JOURNAL	Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center Research Institute, Carcinogenesis Division; Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@gan2.ncc.go.jp, Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)			
FEATURES	Location/Qualifiers			
source	1..4500			
	/organism="Rattus norvegicus"			
	/mol_type="mRNA"			
	/strain="WTC"			
	/db_xref="taxon:10116"			
gene	1..4500			
	/gene="Atrn"			
CDS	61..3888			
	/gene="Atrn"			
	/codon_start=1			
	/product="attractin"			
	/protein_id="BAB21018.1"			
	/db_xref="GI:12275312"			
	/translation="MVAAIAAAEATEARLGGYTTATAAPAGKEROHRPCAATGAWRP WPRAGLCPVLSAPLPPPLLPILLFSLILLPLPREAERAAVAAGSGSAAEAK ECDAPVNGRCNPGTQCVCPTGVEQCCHGGRFLTLTGSSGFTVDGPNKYKTK LTLLEGQKIMLRFNHFNATECSNDHLIYIDGDSIYAFLIAAFSGLIIVFGRNET APEVTIYGYALLHFFDAAYNLGFTNITFDMCNPCNGRGECKSSNSTVECEB SENWGESCDIPHTDNCGRPHRGINASPTGSCFPFHQPGPCSIPVPAQSWTR BEYSDLKLPRASHKAENVNGFIMVVGGMFNHSDYSMLAYDLASREMLSNHSVNSV VRYVTHSIALHKDIYMGYGGKIDSTGNVTNLEVFHINESVLLTPKAKDOYAVVGH SAHIVLSSGSRVVMVLFGHCPGLYISVVOEYDLEKNTWSILQTQGLVQGGYGHSS VIDHRTKALYHGKYKAFSAKRYLADLLYRYHVDYQMTILKDSRFFRYLHTATVVS GTMLVFGGNTNDTSMHGAKCFSSDFNADYIACDRWSVLPFRELHVDNRRFGHSAVL HNSTVYFGGNSLLLDVILFTSEQCDHRSEAAACVAGPGIRCLMDTQSSRCTSWE LATEQEAKLSECFRTLDHDCDQDTDCYCTANTNDCHWNCNDHCPVNHSCSTEG QISTAKYDNCPCKPMYCNKTSCTSCALDQCNQWEPNOCIALPENICIGIWHLV GNSCLKITAYENYDNALICRNHNAFLASLTQKVEFVLKOLRLMQSQTSLKLTL TPNVGLRKINYSYCWEDMSPFTNLSLOWPSPSPDAGFCGILSESTGLKAATCIN PLNSVCERPAHSAKOCRFPCALRTACGCTSSSEBECMWSNMKOCVSNAYVASFP FGQCNWYTMSPSPCPCNGCYCTCHCLEQPGCGWCTDPNTNGKGCIGESYKGPVKM PSHASTGNVYPOPLINSMLCLEDSEYNNSP IHCPACQNGSKINQISCERKEDLT GRKHECTISGYPGPTNGKQCPCKNGHASLQNTNTGKFCFTTKGVGEEQOLCEVE NRYOQNPLKGTCCYVTLIDYQFTFSLQEDDRVYTAINFVATDEQNRDLDFINASK NFNLNITWATSFPAQTGTGEVPPVVSXTNINKYKDSFSEKEDFRNHPNITTFVYVSN FTWFIKIQVRVTS"			
Query Match	1.4%	Score 62;	DB 10;	Length 4500;
Best Local	Similarity 100.0%;	Pred. No. 4.6e-21;		
Matches	62;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
ORIGIN				



189283 bp DNA linear ROD 20-AUG-2002  
sequence from clone RP23-228H14 on chromosome 2, complete

REPORT NUMBER	SEGMENT	SOURCE	ORGANISM
1	1	1	1
2	2	2	2
3	3	3	3
4	4	4	4
5	5	5	5
6	6	6	6
7	7	7	7
8	8	8	8
9	9	9	9
10	10	10	10
11	11	11	11
12	12	12	12
13	13	13	13
14	14	14	14
15	15	15	15
16	16	16	16
17	17	17	17
18	18	18	18
19	19	19	19
20	20	20	20
21	21	21	21
22	22	22	22
23	23	23	23
24	24	24	24
25	25	25	25
26	26	26	26
27	27	27	27
28	28	28	28
29	29	29	29
30	30	30	30
31	31	31	31
32	32	32	32
33	33	33	33
34	34	34	34
35	35	35	35
36	36	36	36
37	37	37	37
38	38	38	38
39	39	39	39
40	40	40	40
41	41	41	41
42	42	42	42
43	43	43	43
44	44	44	44
45	45	45	45
46	46	46	46
47	47	47	47
48	48	48	48
49	49	49	49
50	50	50	50
51	51	51	51
52	52	52	52
53	53	53	53
54	54	54	54
55	55	55	55
56	56	56	56
57	57	57	57
58	58	58	58
59	59	59	59
60	60	60	60
61	61	61	61
62	62	62	62
63	63	63	63
64	64	64	64
65	65	65	65
66	66	66	66
67	67	67	67
68	68	68	68
69	69	69	69
70	70	70	70
71	71	71	71
72	72	72	72
73	73	73	73
74	74	74	74
75	75	75	75
76	76	76	76
77	77	77	77
78	78	78	78
79	79	79	79
80	80	80	80
81	81	81	81
82	82	82	82
83	83	83	83
84	84	84	84
85	85	85	85
86	86	86	86
87	87	87	87
88	88	88	88
89	89	89	89
90	90	90	90
91	91	91	91
92	92	92	92
93	93	93	93
94	94	94	94
95	95	95	95
96	96	96	96
97	97	97	97
98	98	98	98
99	99	99	99
100	100	100	100

18. of 27.  
Homo sapiens (human)  
Homo sapiens  
Ginkgo yota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE 1 (bases 1 to 966)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 966)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA
COMMENT On Nov 24, 2000 this sequence version replaced gi:8118072.
FEATURES
source
Location/Qualifiers
1..966
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
49..101
/gene="ATRN"
/number=21
exon
Query Match 1.2%; Score 53; DB 9; Length 966;
Best Local Similarity 100.0%; Pred. No. 4.1e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3417 ATGTGAGTGAATAATCATACCAAGGAACCCCTCTCAGGACATGTTATT 3469
DB 49 ATGTGAGTGAATAATCATACCAAGGAACCCCTCTCAGGACATGTTATT 101
RESULT 65
AR164814 1051 bp DNA linear PAT 17-OCT-2001
LOCUS Sequence 10 from patent US 6274339.
DEFINITION AR164814
ACCESSION AR164814.1 GI:16238023
VERSION AR164814.1
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1051)
AUTHORS Moore, K. and Nagle, D. Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body
weight disorders, including obesity
JOURNAL Patent: US 6274339-A 10 14-AUG-2001;
FEATURES
source
Location/Qualifiers
1..1051
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 1.2%; Score 53; DB 6; Length 1051;
Best Local Similarity 100.0%; Pred. No. 4e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 367 GGCTGGTGGGGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAACTGG 419
DB 571 GGCTGGTGGGGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAACTGG 623
RESULT 66
AB062914S3 1463 bp DNA linear ROD 16-JAN-2002
LOCUS Mesocricetus auratus Atrn gene for attractin, exon 25, 26.
DEFINITION AB062914S3
ACCESSION AB062916
VERSION AB062916.1 GI:16930138

```

```

KEYWORDS 3 of 5
SEGMENT Mesocricetus auratus (golden hamster)
SOURCE Mesocricetus auratus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Mesocricetus.
REFERENCE 1
AUTHORS Kuramoto, T., Nomoto, T., Fujiwara, A., Mizutani, M., Sugimura, T. and
Ushijima, T.
TITLE Insertional mutation of the Attractin gene in the black tremor
hamster
JOURNAL Mamm. Genome 13 (1), 36-40 (2002)
MEDLINE 21635551
PUBMED 11773967
REFERENCE 2 (bases 1 to 1463)
AUTHORS Kuramoto, T. and Ushijima, T.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center
Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,
Chuo-Ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@ncc.go.jp,
URL:http://www.ncc.go.jp/research/rat-genome/, Tel:81-3-3542-2511,
Fax:81-3-5565-1753)
FEATURES
source
Location/Qualifiers
1..1463
/organism="Mesocricetus auratus"
/mol_type="genomic DNA"
/db_xref="taxon:10036"
426..487
/gene="Atrn"
/product="attractin"
/note="CDS is reported in Acc# AB062918"
/number=25
repeat_region 604..663
/rpt_type=tandem
/rpt_unit="ga"
766..844
/gene="Atrn"
/product="attractin"
/note="CDS is reported in Acc# AB062918"
/number=26
exon
ORIGIN
Query Match 1.2%; Score 50; DB 10; Length 1463;
Best Local Similarity 100.0%; Pred. No. 1.6e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3814 CACCACAGCAATTTATGGACCTGGTACAGTTCTTCGACTTCTTCAG 3863
DB 438 CACCACAGCAATTTATGGACCTGGTACAGTTCTTCGACTTCTTCAG 487
RESULT 67
H009384S19 1781 bp DNA linear PRI 14-DEC-2000
LOCUS Homo sapiens attractin precursor (ATRN) gene, exon 22.
DEFINITION H009384S19
ACCESSION AF218907
VERSION AF218907.1 GI:8118073
KEYWORDS
SEGMENT 19 of 27
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1781)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.
and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE 2 (bases 1 to 1781)

```



```

AUTHORS      Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,
              Schlossman, S.F. and Duke-Cohan, J.S.
TITLE        Direct Submission
JOURNAL      Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
              Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
              USA
FEATURES     Location/Qualifiers
             1..1781
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /chromosome="20"
             /map="20p13"
             802..896
             /gene="ATR"
             /number=22
ORIGIN
Query Match      1.2%; Score 50; DB 9; Length 1781;
Best Local Similarity 100.0%; Pred. No. 1.6e-14;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3515 AAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCTGACGAA 3564
Db 847 AAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCTGACGAA 896
RESULT 68
AF301008
LOCUS      Bos taurus attractin mRNA, partial cds.
DEFINITION
ACCESSION AF301008
VERSION   AF301008.1 GI:10304438
KEYWORDS
SOURCE    Bos taurus (cow)
ORGANISM  Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
REFERENCE  1 (bases 1 to 315)
AUTHORS   Glad, T.W., Granholm, N.H., Westby, C.A., Marshall, D.M. and Kraft, E.
TITLE     Bovine mahogany/attractin cDNA, including exons 26-29
JOURNAL   Unpublished
REFERENCE  2 (bases 1 to 315)
AUTHORS   Glad, T.W., Granholm, N.H., Westby, C.A., Marshall, D.M. and Kraft, E.
TITLE     Direct Submission
JOURNAL   Submitted (29-AUG-2000) Bio/Micro, South Dakota State University,
              2140D, 252 NPB, Brookings, SD 57007, USA
FEATURES     Location/Qualifiers
             1..315
             /organism="Bos taurus"
             /mol_type="mRNA"
             /db_xref="taxon:9913"
             /tissue_type="cerebrum"
             <1..>315
             /note="mahogany"
             /codon_start=2
             /product="attractin"
             /protein_id="AAG16254.1"
             /db_xref="GI:10304439"
             /translation="HSNFMDLVQFPVTFVFFSCFLSLLLVAAVWVKIKSCWASRRRRL
             LREMQMASRFPASVNVVALETDEPDLIGSINKVPKPIALEPCFGNKAIVLNVFR
             LPR"
CDS
Query Match      1.1%; Score 49; DB 4; Length 315;
Best Local Similarity 100.0%; Pred. No. 6.8e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3878 TGCTCCTGGTGGCTGCTGCTGTTGGAGATCAACAAGTTGTTGGC 3926
Db 63 TGCTCCTGGTGGCTGCTGCTGTTGGAGATCAACAAGTTGTTGGC 111

```

```

RESULT 69
AB062914S2
LOCUS      Mesocricetus auratus Atrn gene for attractin, exon 23, 24.
DEFINITION
ACCESSION AB062915
VERSION   AB062915.1 GI:16930137
KEYWORDS
SOURCE    Mesocricetus auratus (golden hamster)
ORGANISM  Mesocricetus auratus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Mesocricetus.
REFERENCE  1
AUTHORS   Kuramoto, T., Nomoto, T., Fujiwara, A., Mizutani, M., Sugimura, T. and
              Ushijima, T.
TITLE     Insertional mutation of the Attractin gene in the black tremor
              hamster
JOURNAL   Mamm. Genome 13 (1), 36-40 (2002)
MEDLINE   21695551
PubMed    11773967
REFERENCE  2 (bases 1 to 4862)
AUTHORS   Kuramoto, T. and Ushijima, T.
TITLE     Direct Submission
JOURNAL   Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center
              Research Institute, Carcinogenesis Division, Tsukiji 5-1-1,
              Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramo@ncc.go.jp,
              URL:http://www.ncc.go.jp/research/rat-genome/, Tel:81-3-3542-2511,
              Fax:81-3-5565-1753)
FEATURES     Location/Qualifiers
             1..4862
             /organism="Mesocricetus auratus"
             /mol_type="genomic DNA"
             /db_xref="taxon:10036"
             594..672
             /gene="Atrn"
             /product="attractin"
             /note="CDS is reported in Acc# AB062918"
             /number=23
             2976..3133
             /gene="Atrn"
             /product="attractin"
             /note="CDS is reported in Acc# AB062918"
             /number=24
             3782..3886
             /gene="Atrn"
             3782..3886
             /rpt_type=tandem
             /rpt_unit="aag"
ORIGIN
Query Match      1.1%; Score 48; DB 10; Length 4862;
Best Local Similarity 100.0%; Pred. No. 1.7e-13;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3754 ACTTCTTTGTTATGTCAGTAATTTACCTGGCCCAAAATTCAG 3801
Db 3086 ACTTCTTTGTTATGTCAGTAATTTACCTGGCCCAAAATTCAG 3133
RESULT 70
E26747
LOCUS      E26747
DEFINITION Structure and function of novel biorhythm marker gene.
ACCESSION E26747
VERSION   E26747.1 GI:13026325
KEYWORDS JP 1999169185-A/1.
SOURCE    Rattus sp.
ORGANISM  Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

REFERENCE 1 (bases 1 to 1012)  
AUTHORS Mario,I., Yutaka,S., Satoru,S. and Mikio,K.  
TITLE Structure and function of novel biorhythm marker gene  
JOURNAL Patent: JP 1999169185-A 1 29-JUN-1999;  
AGENCY OF IND SCIENCE & TECHNOL,OTSUKA PHARMACEUT CO LTD  
COMMENT OS Rattus sp. (Wistar rat)  
PN JP 1999169185-A/1  
PF 29-JUN-1999  
PD 12-DEC-1997 JP 1997362890  
PR MARIO ISHIDA, YUTAKA SADAKANE, SATORU SUZUKI, MIKIO KIKUCHI PC  
C12N15/09, C07K14/47, C12Q1/68//C12N5/10, C12P21/02, (C12N15/09, PC  
C12R1.91),  
PC (C12N5/10, C12R1.91), (C12P21/02, C12R1.91), (C12N15/00, C12N5/00,  
C12N15/00, C12R1.91), (C12N5/00, C12R1.91)  
CC Strandedness: Single;  
CC Topology: Linear;  
FH Key 1. .720.  
FT CDS Location/Qualifiers  
FEATURES  
source  
1..1012  
/organism="Rattus sp."  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10118"  
ORIGIN  
Query Match 1.1%; Score 47; DB 6; Length 1012;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2698 AGTACTCGGGAGCTGAAGCTGCAACCTGCATCAACCCACTCAATGG 2744  
DB 466 AGTACTCGGGAGCTGAAGCTGCAACCTGCATCAACCCACTCAATGG 512  
RESULT 71  
AB069689S15  
LOCUS Rattus norvegicus 1207 bp DNA linear ROD 04-APR-2002  
DEFINITION Rattus norvegicus Atrn gene for membrane attractin, exon 16.  
ACCESSION AB049236  
VERSION AB049236.1 GI:12275377  
KEYWORDS  
SEGMENT  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM  
REFERENCE  
AUTHORS Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,  
Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.,  
and Serikawa,T.  
TITLE Attractin/mahogany/zitter plays a critical role in myelination of  
the central nervous system  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)  
MEDLINE 2143347  
PUBMED 11209055  
REFERENCE 2 (bases 1 to 1207)  
AUTHORS Kuramoto,T. and Ushijima,T.  
TITLE Direct Submission  
JOURNAL Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center  
Research Institute, Carcinogenesis Division; Teukiji 5-1-1,  
Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@gan2.ncc.go.jp,  
Tel:81-33542-2511, Fax:81-5565-1753)  
LOCATION/Qualifiers  
FEATURES  
source  
1..1207  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/chromosome="3"  
488..707  
/gene="Atrn"  
/product="membrane attractin"  
exon

ORIGIN  
Query Match 1.1%; Score 47; DB 10; Length 1207;  
Best Local Similarity 100.0%; Pred. No. 7e-13;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2698 AGTACTCGGGAGCTGAAGCTGCAACCTGCATCAACCCACTCAATGG 2744  
DB 641 AGTACTCGGGAGCTGAAGCTGCAACCTGCATCAACCCACTCAATGG 687  
RESULT 72  
AR164813  
LOCUS Sequence 8 from patent US 6274339. linear PAT 17-OCT-2001  
DEFINITION AR164813  
ACCESSION AR164813  
VERSION AR164813.1 GI:16238022  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2419)  
AUTHORS Moore,K. and Magle,D.Lynn.  
TITLE Methods and compositions for the diagnosis and treatment of body  
weight disorders, including obesity  
JOURNAL Patent: US 6274339-A 8 14-AUG-2001;  
FEATURES  
source  
1..2419  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 1.1%; Score 47; DB 6; Length 2419;  
Best Local Similarity 100.0%; Pred. No. 6.5e-13;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1624 CAGATGTGACCACTTCTTAAGGACAGCCGATTTTCGGTACTTGA 1670  
DB 1890 CAGATGTGACCACTTCTTAAGGACAGCCGATTTTCGGTACTTGA 1936  
RESULT 73  
AL954766\_2  
WPCOMMENT  
Sequence split into 4 fragments LOCUS AL954766 Accession AL954766  
Fragment Name Begin End  
AL954766\_0 1 110000  
AL954766\_1 100001 210000  
AL954766\_2 200001 310000  
AL954766\_3 300001 356753  
Continuation (3 of 4) of AL954766 from base 200001 (AL954766 Mus musculus chromosome X c  
Query Match 1.1%; Score 47; DB 2; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 4.1e-13;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3829 ATGGACCTGGTACAGTTCTTCGTGACTTCTTCAGTTGTTTCCTCTC 3875  
DB 60794 ATGGACCTGGTACAGTTCTTCGTGACTTCTTCAGTTGTTTCCTCTC 60840  
RESULT 74  
BX004762\_0/c  
WPCOMMENT  
Sequence split into 5 fragments LOCUS BX004762 Accession BX004762  
Fragment Name Begin End  
BX004762\_0 1 110000  
BX004762\_1 100001 210000  
BX004762\_2 200001 310000  
BX004762\_3 300001 410000  
BX004762\_4 400001 442165

LOCUS BX004762 442165 bp DNA linear HTG 12-APR-2003  
DEFINITION Mus musculus chromosome X clone RP23-146C4, WORKING DRAFT SEQUENCE,  
4 unordered pieces.  
ACCESSION BX004762  
VERSION BX004762.6 GI:29824756  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_ACTIVEFIN; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Chapman, J.  
1 (bases 1 to 442165)  
REFERENCE Direct Submission  
AUTHORS Submitted (11-APR-2003) Wellcome Trust Sanger Institute, Hinxton,  
TITLE Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
JOURNAL humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
COMMENT On Apr 13, 2003 this sequence version replaced gi:29125168.  
Sequence from the Mouse Genome Sequencing Consortium whole genome  
shotgun may have been used to confirm this sequence. Sequence data  
from the whole genome shotgun alone has only been used where it has  
a paired quality of at least 30.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
----- Project Information  
Center project name: BM146C4  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 229053 bases at least Q40  
Consensus quality: 229393 bases at least Q30  
Consensus quality: 229535 bases at least Q20  
Insert size: 441865; sum-of-contigs  
Insert size: 188914; 6.6% error; agarose-fp  
Quality coverage: 3.34x in Q20 bases; sum-of-contigs Quality  
coverage: 7.96x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 48919: contig of 48919 bp in length  
\* 48920 49019: gap of 100 bp  
\* 49020 130781: contig of 81762 bp in length  
\* 130782 130881: gap of 100 bp  
\* 130882 229564: contig of 95083 bp in length  
\* 229565 230064: gap of 100 bp  
\* 230065 442165: contig of 21201 bp in length.  
\*  
\* Location/Qualifiers  
\* 1..442165  
\* /organism="Mus musculus"  
\* /mol\_type="genomic DNA"  
\* /db\_xref="taxon:10090"  
\* /chromosome="X"  
\* /clone="RP23-146C4"  
\* /clone\_lib="RPCI-23"  
\* 1..48919  
\* /note="assembly fragment:00147"  
\* fragment\_chain:1  
\* 49020..130781  
\* /note="assembly fragment:00500"  
\* fragment\_chain:1  
\* 130882..229564  
\* /note="assembly fragment:00870"  
\* fragment\_chain:1  
\* 230065..442165  
\* /note="assembly fragment:02147"

ORIGIN  
Query Match 1.1%; Score 47; DB 2: Length 110000;  
Best Local Similarity 100.0%; Pred.No. 4.1e-13;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 3829 ATGGACCTGGTACAGTCTTCGTCGACTTCTTCAGTGTGTTCTCTC 3875  
Db 101301 ATGGACCTGGTACAGTCTTCGTCGACTTCTTCAGTGTGTTCTCTC 101255  
RESULT 75  
BX004762 1/c  
WPCOMMENT  
Sequence split into 5 fragments LOCUS BX004762 Accession BX004762  
Fragment Name Begin End  
BX004762\_0 1 110000  
BX004762\_1 100001 210000  
BX004762\_2 200001 310000  
BX004762\_3 300001 410000  
BX004762\_4 400001 442165  
Continuation (2 of 5) of BX004762 from base 100001 (BX004762 Mus musculus chromosome X c  
Query Match 1.1%; Score 47; DB 2: Length 110000;  
Best Local Similarity 100.0%; Pred.No. 4.1e-13;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 3829 ATGGACCTGGTACAGTCTTCGTCGACTTCTTCAGTGTGTTCTCTC 3875  
Db 1301 ATGGACCTGGTACAGTCTTCGTCGACTTCTTCAGTGTGTTCTCTC 1255  
Search completed: March 1, 2004, 18:06:08  
Job time : 15969 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 1, 2004, 12:56:07 ; Search time 9532 Seconds

(without alignments)  
13439.861 Million cell updates/sec

Title: US-09-787-097-13

Perfect score: 4290

Sequence: 1 atggtggccgacggcgccg.....agccctgggacctgcattctga 4290

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 275113289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

EST.\*

1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_htc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_man.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2368	55.2	3976	29	AY418587
2	808	18.8	1201	13	BX440935
3	681	15.9	681	12	BM783739
4	680	15.9	680	12	BM783258

5	680	15.9	680	12	BM783773
6	680	15.9	680	12	BM783773
7	677	15.8	677	12	BM783718
8	673	15.7	673	12	BM783685
9	632	14.7	632	12	BM773003
10	632	14.7	632	12	BM783264
11	612	14.3	792	14	CD644584
12	581	13.5	683	12	BM783788
13	580	13.5	633	10	BE894518
14	576	13.4	677	12	BM783821
15	556	12.0	917	10	BF672370
16	516	12.0	962	13	EX431593
17	485	11.3	486	12	BM772968
18	471	11.0	591	12	BI850289
19	435	10.1	598	10	BF853006
20	426	9.9	746	14	CB242644
21	425	9.9	595	10	BF853681
22	413	9.6	744	13	EQ183626
23	402	9.4	567	10	BF854487
24	400	9.3	544	10	BF853962
25	399	9.3	643	14	CA305924
26	390	9.1	3671	29	AY418588
27	389	9.1	739	14	CA502816
28	382	8.9	674	12	EG563729
29	369	8.6	502	9	AL048842
30	363	8.5	991	10	BE880905
31	360	8.4	484	10	BF963698
32	355	8.3	953	10	BF313142
33	353	8.2	560	10	BF853965
34	352	8.2	390	12	EG673763
35	347	8.1	681	12	BM981492
36	335	7.8	493	10	BE259981
37	333	7.8	706	12	BI089332
38	326	7.6	484	12	BI004303
39	309	7.2	399	10	BF880955
40	300	7.0	671	9	AI818764
41	299	7.0	383	9	AI372809
42	298	6.9	427	10	AW579586
43	283	6.6	320	9	AI372810
44	283	6.6	400	9	AA350293
45	281	6.6	483	10	BF853021
46	270	6.3	281	10	EG015062
47	270	6.3	396	9	AA350292
48	269	6.3	524	10	BE709733
49	253	5.9	778	13	BU531905
50	249	5.8	589	9	AI818750
51	243	5.7	313	10	BF928756
52	241	5.6	416	14	BF7660
53	225	5.2	445	10	BE671400
54	223	5.2	464	9	AI076479
55	185	4.3	500	13	EX644466
56	176	4.1	517	14	T69367
57	171	4.0	609	10	BF206327
58	167	3.9	297	9	AA368140
59	165	3.8	265	13	BQ370569
60	165	3.8	980	13	BU158314
61	136	3.2	355	14	T92696
62	135	3.1	413	10	BF475948
63	129	3.0	279	9	AA297214
64	126	2.9	370	9	AA771958
65	125	2.9	319	9	AA298732
66	123	2.9	230	9	AA297697
67	123	2.9	683	10	BF700035
68	119	2.8	837	28	BZ610572
69	115	2.7	134	12	BG945444
70	113	2.6	530	13	EX669805
71	110	2.6	363	29	CE218575
72	106	2.5	942	10	BF316716
73	102	2.4	496	13	BQ307981
74	97	2.3	160	10	BE171376
75	94	2.2	517	28	AQ826412
76	92	2.1	527	14	HI1351
77	89	2.1	644	14	CB425680

BM783773	K-EST0061
EG678679	602624477
BM783718	K-EST0061
BM783685	K-EST0061
BM773003	K-EST0057
BM783264	K-EST0061
CD644584	AGENCOURT
BM783788	K-EST0061
BE894518	K-EST0061
BM783821	K-EST0061
BF672370	602150605
EX431593	EX431593
BM772968	K-EST0057
BI850289	imageoc 1
BF853006	MR2-ENO09
CB242644	UI-CF-FNO
BF853681	MR2-ENO09
EQ183626	UI-H-EJ0-
BF854487	MR2-ENO09
BF853962	MR2-ENO09
CA305924	UI-H-PT1-
AY418588	Par tlogl
CA502816	UI-CF-FNO
EG563729	602584512
AL048842	DKF2P434G
BE880905	601490448
BF963698	PM4-NN120
BF313142	601897661
BF853965	MR2-ENO09
EG673763	9010 cll1
BM981492	UI-CF-EN1
BE259981	601148539
BI089332	602853292
BI004303	PMO-HNO07
BF880955	OVI-ET018
AI818764	willie04.x
AI372809	EST175353
AW579586	MRO-HT020
AI372810	EST175354
AA350293	EST17512
BF853021	MR2-ENO09
EG015062	OV2-GNO20
AA350292	EST17511
BF709733	IL3-HT061
BU531905	AGENCOURT
AI818750	willie03.x
BF928756	IL5-NT022
BF7660	VP89901.s1
BE671400	7es0d11.x
AI076479	oz28804.x
EX644466	DKF2P781N
T69367	YC37902.r1
BF206327	601869664
AA368140	EST179344
BQ370569	CM4-GN033
BU158314	AGENCOURT
T92696	ye26f05.r1
BF475948	nao42h02.
AA297214	EST112757
AA771958	ai37e06.s8
AA298732	EST114356
AA297697	EST113254
BF700035	602127513
BZ610572	WHAAD221R
BG945444	PM2-AN009
EX669805	EX669805
CE218575	tigr-gss-
BF316716	601903293
BQ307981	MRO-BT550
BE171376	RC1-HT054
AQ826412	HS_504-B
HI1351	Yml13502.r1
CB425680	600762 MA

```

78      86      2.0      693      9      AV725515
79      325      1.8      10      BF773440
80      76      1.8      287      10      BF804241
81      76      1.8      493      10      BE770345
82      76      1.8      501      10      BE770333
83      72      1.7      450      13      BQ373587
84      67      1.6      532      13      BX103722
85      65      1.5      529      12      BM537406
86      65      1.5      694      29      CE775705
87      62      1.4      290      13      BY483137
88      62      1.4      399      10      AW124112
89      62      1.4      429      9      AI575016
90      62      1.4      441      9      AA901110
91      62      1.4      457      13      BY483004
92      62      1.4      551      10      BE111505
93      62      1.4      565      10      BE112266
94      62      1.4      714      10      BE650220
95      62      1.4      735      13      BY761390
96      62      1.4      770      13      BY761276
97      62      1.4      3976      29      AY418589
98      61      1.4      312      9      AI372811
99      61      1.4      610      13      BQ183082
100     58      1.4      316      14      R84298

```

## ALIGNMENTS

```

RESULT 1
LOCUS      AY418587
DEFINITION Homo sapiens ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
            3976 bp DNA linear GSS 17-DEC-2003
ACCESSION  AY418587
VERSION     AY418587.1 GI:39774547
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 3976)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
            Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL     Science 302 (5652), 1960-1963 (2003)
PUBMED     14671302
REFERENCE   2 (bases 1 to 3976)
AUTHORS     Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
            Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
            Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
            Direct Submission
            Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
            These sequences were made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES    Location/Qualifiers
            source          1..3976
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
            gene            <1..>3976
                        /gene="ATRN"
                        /locus_tag="HCM6607"
ORIGIN
Query Match      55.2%; Score 2368; DB 29; Length 3976;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2418; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1872  CACCATGTAAGTGTTCGGTGGTTTCAATAGTCTCTCTCTCAGCGACATCTCTGGTATTTCAC 1931
DB      1558  CACCATGTAAGTGTTCGGTGGTTTCAATAGTCTCTCTCTCAGCGACATCTCTGGTATTTCAC 1617
QY      1932  CTCGGAACAAGTGTGATGCGCATCGGAGTGAAGCGCTTGTATTAGCAGCAGGACCTGGTAT 1991
DB      1618  CTCGGAACAAGTGTGATGCGCATCGGAGTGAAGCGCTTGTATTAGCAGCAGGACCTGGTAT 1677
QY      1992  TCGGTGTGTGTGGAAACACAGGTCCTCTCAGTGTATCTCTGTGGCGCTGGCAACTGATGA 2051
DB      1678  TCGGTGTGTGTGGAAACACAGGTCCTCTCAGTGTATCTCTGTGGCGCTGGCAACTGATGA 1737
QY      2052  ACAAGAGAAAGTTAAATCAGATGTTTCCAAAAGAACTCTTGACCATGACAGATG 2111
DB      1738  ACAAGAGAAAGTTAAATCAGATGTTTCCAAAAGAACTCTTGACCATGACAGATG 1797
QY      2112  TGACCAGACACAGAGATTGTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTGCAA 2171
DB      1798  TGACCAGACACAGAGATTGTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTGCAA 1857
QY      2172  TGACCAATTGTGTCCAGGAACACAGCTGTCTCAGAGCCAGATCTCCATTTTAGTGA 2231
DB      1858  TGACCAATTGTGTCCAGGAACACAGCTGTCTCAGAGCCAGATCTCCATTTTAGTGA 1917
QY      2232  TGAGAAATTGCCCCAAGGATAACCCCATGTAACCTGTACTGTAAACAAGAGACAGCTGCAGGAG 2291
DB      1918  TGAGAAATTGCCCCAAGGATAACCCCATGTAACCTGTACTGTAAACAAGAGACAGCTGCAGGAG 1977
QY      2292  CTGTGCCCTGGACACGAACTGCCAGTGGGAGCCCGGAATCAGGAGTGCATGCCCTGCC 2351
DB      1978  CTGTGCCCTGGACACGAACTGCCAGTGGGAGCCCGGAATCAGGAGTGCATGCCCTGCC 2037
QY      2352  CGAAATAATCTGTGGCATTGGCTGTCATTTGGTTGGAACTCATGTTTGAATAATTACTAC 2411
DB      2038  CGAAATAATCTGTGGCATTGGCTGTCATTTGGTTGGAACTCATGTTTGAATAATTACTAC 2097
QY      2412  TGCCAAAGGAGAAATTATGACAATGCTTAAATTTGTTGTAGGAACCAACAATGCCCTTTGGC 2471
DB      2098  TGCCAAAGGAGAAATTATGACAATGCTTAAATTTGTTGTAGGAACCAACAATGCCCTTTGGC 2157
QY      2472  TTCTCTTACACCCAGAGAAGAGTAGATTGTCTTTAAGCAGCTGCGAATAATGCGAGTC 2531
DB      2158  TTCTCTTACACCCAGAGAAGAGTAGATTGTCTTTAAGCAGCTGCGAATAATGCGAGTC 2217
QY      2532  ATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGCTCGGCCCTTCGGAAGATCAATGT 2591
DB      2218  ATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGCTCGGCCCTTCGGAAGATCAATGT 2277
QY      2592  GTCTCTACTGGTGTGGGAAGATATGTCCTCCCAATTTACAAATAGTTTACTACAGTGGATGCC 2651
DB      2278  GTCTCTACTGGTGTGGGAAGATATGTCCTCCCAATTTACAAATAGTTTACTACAGTGGATGCC 2337
QY      2652  GTCTCAGCCCTCAGTGTGATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGACT 2711
DB      2338  GTCTCAGCCCTCAGTGTGATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGACT 2397
QY      2712  GAAGGCTGCAACTGTCATCAACCCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACCA 2771
DB      2398  GAAGGCTGCAACTGTCATCAACCCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACCA 2457
QY      2772  CAGTGTAAAGCAGTGCAGGACACCATGTGCTTGAGGACAGCATGTGGAGATTGCACCAG 2831
DB      2458  CAGTGTAAAGCAGTGCAGGACACCATGTGCTTGAGGACAGCATGTGGAGATTGCACCAG 2517
QY      2832  CGGCAGCTCTCAGTGTGATGTGTGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCTA 2891
DB      2518  CGGCAGCTCTCAGTGTGATGTGTGTGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCTA 2577
QY      2892  TGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2951
DB      2578  TGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2637
QY      2952  TGAAAATTGTTTTCAGGCTACTGTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 3011

```

2638 TGAATAATTGTTAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTG 2697  
3012 GTGTACTGATCCAGCAATCTGCTGCAAGGAAATGATAGAGGTTCTTATTAAGGACC 3071  
2698 GTGTACTGATCCAGCAATCTGCTGCAAGGAAATGATAGAGGTTCTTATTAAGGACC 2757  
3072 AGTGAAGATGCTTTCGCAAGCCCTCTACAGGAAATTTCTATCCAGCCCTGCTCAATTG 3131  
2758 AGTGAAGATGCTTTCGCAAGCCCTCTACAGGAAATTTCTATCCAGCCCTGCTCAATTG 2817  
3132 CAGCATGTCTAGAGACAGCAGATACACTGCTCTTTTCAATTCATCTGCTGCTGCTGCA 3191  
2818 CAGCATGTCTAGAGACAGCAGATACACTGCTCTTTTCAATTCATCTGCTGCTGCTGCA 2877  
3192 ATGCAAGCGGCACAGTAATGCTCAATCAGAGCATCTGTGAGAAGTGTGAGAAGCTGAC 3251  
2878 ATGCAAGCGGCACAGTAATGCTCAATCAGAGCATCTGTGAGAAGTGTGAGAAGCTGAC 2937  
3252 CACAGGCAAGCATGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCAATGAGG 3311  
2938 CACAGGCAAGCATGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCAATGAGG 2997  
3312 GAAATGTCAGCATGCAAGTGCATGCAATGCGGCTCTCTGCAACACCAACACGGGCAA 3371  
2998 GAAATGTCAGCATGCAAGTGCATGCAATGCGGCTCTCTGCAACACCAACACGGGCAA 3057  
3372 GTGCTTCTGCACCAACCAAGGCGCTCAAGGGGGAGCGAGTCCAGCTATGTGAGTAGAATA 3431  
3058 GTGCTTCTGCACCAACCAAGGCGCTCAAGGGGGAGCGAGTCCAGCTATGTGAGTAGAATA 3117  
3432 TCGATACAGGAAGACCTCTCAGAGGAACATGTTATATATCTTCTTATTCACATCA 3491  
3118 TCGATACAGGAAGACCTCTCAGAGGAACATGTTATATATCTTCTTATTCACATCA 3177  
3492 GTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGGC 3551  
3178 GTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGGC 3237  
3552 TACTCTGAGCAACAAACAGGGATTTGGACATGTTTCAATGCTGCTTCAAGATTTCAA 3611  
3238 TACTCTGAGCAACAAACAGGGATTTGGACATGTTTCAATGCTTCAAGATTTCAA 3297  
3612 CTTCAACATCACCTGGCTGCTGCTTCTCAGCTGGAACCCAGCTGGAAGAGATGCC 3671  
3298 CTTCAACATCACCTGGCTGCTGCTTCTCAGCTGGAACCCAGCTGGAAGAGATGCC 3357  
3672 TGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTGA 3731  
3358 TGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTGA 3417  
3732 TTTTGGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTCACTGGCCCAT 3791  
3418 TTTTGGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTCACTGGCCCAT 3477  
3792 CAAATTCAGATTCGCTTCTCTCAGCAGCAAAATTTATGACCTGGTACAGTTCTTCGT 3851  
3478 CAAATTCAGATTCGCTTCTCTCAGCAGCAAAATTTATGACCTGGTACAGTTCTTCGT 3537  
3852 GACTTCTCAGTTGTTTCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3911  
3538 GACTTCTCAGTTGTTTCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3597  
3912 ACAAGTTGTTGGGCTCCAGAGCTAGAGACCAATCTTCTCAGAGATGCAACAGATGGC 3971  
3598 ACAAGTTGTTGGGCTCCAGAGCTAGAGACCAATCTTCTCAGAGATGCAACAGATGGC 3657  
3972 CAGCGCTCCCTTTCCTCTGTAATGTCGCTTGGAAACAGATGAGGAGCTCTGATCT 4031  
3658 CAGCGCTCCCTTTCCTCTGTAATGTCGCTTGGAAACAGATGAGGAGCTCTGATCT 3717  
4032 TATTGGGGGAGTATAAGATCTGTTCCCAACCAATTCACATGCAATGCTTAATTTCTTAGGAACA 4091

3718 TATTGGGGGAGTATAAGATCTGTTCCAAACCCANTTGCATCTGGAGCGTGTGTTGGCAA 3777  
4092 CAAAGCCGCTGTCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCTTGGTGGCATCCCTCC 4151  
3778 CAAAGCCGCTGTCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCTTGGTGGCATCCCTCC 3837  
4152 TCCTGGGAGTCAGATCTTCTGCTGGCCAGGCTTGTGAGCATTTCTCAGCAGATGCC 4211  
3838 TCCTGGGAGTCAGATCTTCTGCTGGCCAGGCTTGTGAGCATTTCTCAGCAGATGCC 3897  
4212 GATAGTGTACAGGAGAAAGTCAGAGGCTGTGAGAAACCGAAGCAGCAGCCCTTCGACA 4271  
3898 GATAGTGTACAGGAGAAAGTCAGAGGCTGTGAGAAACCGAAGCAGCAGCCCTTCGACA 3957  
4272 GCCTGGGACCTGCATCTGA 4290  
3958 GCCTGGGACCTGCATCTGA 3976

RESULT 2  
BX440935 1201 bp mRNA linear EST 15-MAY-2003  
BX440935 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone  
CS0DF012YL19 5-PRIME, mRNA sequence.  
ACCESSION BX440935  
VERSION BX440935.1 GI:30781889  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1. (bases 1 to 1201)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
Bp 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 10212.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0DF012CF10QPl&cluster=10212.f. Contact :  
Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS0DF012CF10QPl.

FEATURES  
source  
1. .1201  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DF012YL19"  
/tissue\_type="FETAL BRAIN"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL BRAIN"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-strand cDNA was digested with Not I and  
cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
vector. Library was not normalized."

ORIGIN  
Query Match 18.8%; Score 808; DB 13; Length 1201;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 908; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2337 GTGCATTCGCTGCGGCAAAATATCTGTGGCATTTGGTGGCAACTCATG 2396  
Db 57 GTGCATTCGCTGCGGCAAAATATCTGTGGCATTTGGTGGCAACTCATG 116  
QY 2397 TTTGAAATTTACTCTGCGCAAGAGAAATATGACATGCTAAATTTCTTAGGAACA 2456

Db 117 TTTGAAATTAATCTACTGCGAAGGAGAAATTATGCAATGCTAAATTTGTTCTGTAGAACCA 176  
Qy 2457 CAATCCCTTTTGGCTTCTCTTACAAACCCAGAGAGTAGAATTTGTCCTTAAGCAGCT 2516  
Db 177 CAATGCCCTTTTGGCTTCTCTTACAAACCCGATGATGAAGGTAGAAATTTGTCCTTAAGCAGCT 236  
Qy 2517 GCGAATATGAGTCATCTCAGAGATGTCACAGTCACCTTAACCCATGGTGGCCT 2576  
Db 237 GCGAATATGAGTCATCTCAGAGATGTCACAGTCACCTTAACCCATGGTGGCCT 296  
Qy 2577 TCGGAAGATCAATGTGCTCTACTGTGCTGGGAGATATGTCCTCCCAATTTACAAATAGTTT 2636  
Db 297 TCGGAAGATCAATGTGCTCTACTGTGCTGGGAGATATGTCCTCCCAATTTACAAATAGTTT 356  
Qy 2637 ACTACAGTGGATGCGCTGTGAGCCAGTCAGTCGTGATGCTGGATTTCTGTGGAATTTTATCAGAAC 2696  
Db 357 ACTACAGTGGATGCGCTGTGAGCCAGTCAGTCGTGATGCTGGATTTCTGTGGAATTTTATCAGAAC 416  
Qy 2697 CAGTACTCGGAGCTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGTGA 2756  
Db 417 CAGTACTCGGAGCTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGTGA 476  
Qy 2757 AAGGCTGCAACACACAGTGTGAAGCAGTCCGACACCACTGTGCTGTGAGCAGCATG 2816  
Db 477 AAGGCTGCAACACACAGTGTGAAGCAGTCCGACACCACTGTGCTGTGAGCAGCATG 536  
Qy 2817 TGGAGATTGCCACGCGCAGCTCTGAGTCATGCTGGTGAGCAACATGAGCAGTGTGT 2876  
Db 537 TGGAGATTGCCACGCGCAGCTCTGAGTCATGCTGGTGAGCAACATGAGCAGTGTGT 596  
Qy 2877 GGAATCCCAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTATACGAT 2936  
Db 597 GGAATCCCAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTATACGAT 656  
Qy 2937 GAGACCTGCCCCCTGAAATTTGTCAGCTACTGTACCTGTAGTCAATGCTTGGAGCA 2996  
Db 657 GAGACCTGCCCCCTGAAATTTGTCAGCTACTGTACCTGTAGTCAATGCTTGGAGCA 716  
Qy 2997 ACCAGCTGTGCTGTGCTGATGCTGATCCAGCAATCTGCGCAAGGAAATGCTAGAGGG 3056  
Db 717 ACCAGCTGTGCTGTGCTGATGCTGATCCAGCAATCTGCGCAAGGAAATGCTAGAGGG 776  
Qy 3057 TTCCTATAAGACACAGTGAAGTGCCTTCCGAGCCCTTACAGGAATTTCTATCCACA 3116  
Db 777 TTCCTATAAGACACAGTGAAGTGCCTTCCGAGCCCTTACAGGAATTTCTATCCACA 836  
Qy 3117 GCCCTGTCTCAATCCAGCATGTCTAGAGCAGCAGATACAACTGGTCTTTTCAATCA 3176  
Db 837 GCCCTGTCTCAATCCAGCATGTCTAGAGCAGCAGATACAACTGGTCTTTTCAATCA 896  
Qy 3177 CTGTCCAGCTTGCCTGCAACGCGCCACAGTAAATGTCATCAATCAGAGCATCTGTGAGAA 3236  
Db 897 CTGTCCAGCTTGCCTGCAACGCGCCACAGTAAATGTCATCAATCAGAGCATCTGTGAGAA 956  
Qy 3237 GTGTGAGAAC 3246  
Db 957 GTGTGAGAAC 966

RESULT 3  
LOCUS BM783739  
DEFINITION K-EST0061730 S5SNU484 Homo sapiens CDNA clone S5SNU484-34-H05 5',  
mRNA sequence.  
ACCESSION BM783739  
VERSION BM783739.1 GI:19131971  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 681)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 34 row: H column: 05  
High quality sequence stop: 681.  
FEATURES  
Location/Qualifiers  
1..681  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S5SNU484-34-H05"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/cell\_line="SNU-484"  
/lab\_host="Top10F"  
/clone\_lib="S5SNU484"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;  
Site 2: NotI; The poly (A) + RNA was decapped with tobacco  
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
including EcoRI site by treatment of T4 RNA ligase. The  
first strand cDNA was synthesized from oligo dt-selected  
mRNA by priming with dt-tailed vector. The dt-tailed  
vector was adjusted to have about 60nt. The cDNA vector  
was circularized with E. coli DNA ligase after digestion  
of EcoRI which site is also included in vector. An RNA  
strand converted to a DNA strand by Okayama-Berg method.  
The obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation  
method."

## ORIGIN

Query Match 15.9%; Score 681; DB 12; Length 681;  
Best/Local Similarity 100.0%; Pred. No. 0;  
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2512 CAGCTGCGAATTAATCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGCTC 2571  
Db 1 CAGCTGCGAATTAATCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGCTC 60  
Qy 2572 GGCCTTCGGAAGATCAATGTGCTCTACTGGTGCTGGGAGATATGTCCTCCCATTTACAAT 2631  
Db 61 GGCCTTCGGAAGATCAATGTGCTCTACTGGTGCTGGGAGATATGTCCTCCCATTTACAAT 120  
Qy 2632 AGTTTACTACAGTCGATGCCGTCTCAGGCCAGTCGTGATGCTGGATTCGTGGAAATTTATCA 2691  
Db 121 AGTTTACTACAGTCGATGCCGTCTCAGGCCAGTCGTGATGCTGGATTCGTGGAAATTTATCA 180  
Qy 2692 GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTG 2751  
Db 181 GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTG 240  
Qy 2752 TGTGAAAGCCCTGCAAAACACAGTCGCTAAGCAGTCCCGGACACCATGTGCTTGAAGACA 2811  
Db 241 TGTGAAAGCCCTGCAAAACACAGTCGCTAAGCAGTCCCGGACACCATGTGCTTGAAGACA 300  
Qy 2812 GCATCTGGAGATTCACCCAGCGGCAGCTCTGAGTCATGTGGTGAGCAACATGAAGCAG 2871  
Db 301 GCATCTGGAGATTCACCCAGCGGCAGCTCTGAGTCATGTGGTGAGCAACATGAAGCAG 360  
Qy 2872 TGTGTGGACTCCAATGCCCTATGTGCGCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2931  
Db 361 TGTGTGGACTCCAATGCCCTATGTGCGCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420  
Qy 2932 ACGATGAGCACCTGCCCCCTCGAAATTTGTCAGGCTACTGTACCTGTAGTCATTTGCTGTG 2991



```

|||||
Db 421 ACGATGAGCAGCTGCCCCCTGAAATTTGTTTCAGGCTACTGTACTGTACTGTCTGTTG 480
Qy 2992 GAGCAACAGAGCTGTGGTGTGTACTGATCCAGCAATACTGCGAAAGGAAATGCATA 3051
Db 481 GAGCAACAGAGCTGTGGTGTGTACTGATCCAGCAATACTGCGAAAGGAAATGCATA 540
Qy 3052 GAGGTTTCTATAAGACACAGTGAAGATGCTTTCGAAAGCCCTACAGAAATTTCTAT 3111
Db 541 GAGGTTTCTATAAGACACAGTGAAGATGCTTTCGAAAGCCCTACAGAAATTTCTAT 600
Qy 3112 CCAGAGCCCTGCTCAATTCAGCATGTGCTAGAGACAGCAGATACAATGCTGTTTC 3171
Db 601 CCAGAGCCCTGCTCAATTCAGCATGTGCTAGAGACAGCAGATACAATGCTGTTTC 660
Qy 3172 ATTCAGTGTCCAGCTTGCCAA 3192
Db 661 ATTCAGTGTCCAGCTTGCCAA 681

```

RESULT 4  
 BM783258  
 LOCUS  
 DEFINITION K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5',  
 21C Frontier Korean EST Project 2001  
 mRNA sequence.

ACCESSION BM783258  
 VERSION BM783258.1 GI:19131490  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 680)  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS

Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470  
 Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr  
 Plate: 31 row: H column: 08

High quality sequence stop: 680.

Location/Qualifiers

FEATURES  
 source

1..680

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S5SNU484-31-H08"

/sex="M"

/tissue\_type="Stomach"

/cell\_type="Epithelial"

/lab\_host="Top10F"

/clone\_lib="S5SNU484"

/note="Organ: Stomach; Vector: pTZ19RP1; Site: 1: EcoRI;  
 Site: 2: NotI; The poly (A)+ RNA was decapped with tobacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 RNA ligase. The  
 first strand cDNA was synthesized from oligo dt-tailed  
 mRNA by priming with dt-tailed vector. The dt-tailed  
 vector was adjusted to have about 60nt. The cDNA vector  
 was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a DNA strand by Okayama-Berg method.  
 The obtained cDNA vectors were used for transfection of  
 competent cells E. coli Top10F by electroporation  
 method."

ORIGIN

Query Match 15.9%; Score 680; DB 12; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-313;  
 Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 2512 CAGCTGCGAATATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC 2571
Db 1 CAGCTGCGAATATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC 60
Qy 2572 GGCCTTCGGAAGATCAATGTCTTACTGTCTGGGAAGATATGTCCCATTTTCAAAAT 2631
Db 61 GGCCTTCGGAAGATCAATGTCTTACTGTCTGGGAAGATATGTCCCATTTTCAAAAT 120
Qy 2632 AGTTTACTACAGTGGATGCCCTGTAGCCAGTGTCTGGATTCGTGGAATTTTATCA 2691
Db 121 AGTTTACTACAGTGGATGCCCTGTAGCCAGTGTCTGGATTCGTGGAATTTTATCA 180
Qy 2692 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGC 2751
Db 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGC 240
Qy 2752 TGTGAAGGCTGCGAAACACACAGTGTAAAGAGTGCCTGCGGACACCATGTGCTTGGAGCA 2811
Db 241 TGTGAAGGCTGCGAAACACACAGTGTAAAGAGTGCCTGCGGACACCATGTGCTTGGAGCA 300
Qy 2812 GCATGTGGAGATTGCAACAGCGGAGCTCTGAGTGCATCTGTGTGAGCAACATGAAGCAG 2871
Db 301 GCATGTGGAGATTGCAACAGCGGAGCTCTGAGTGCATCTGTGTGAGCAACATGAAGCAG 360
Qy 2872 TGTGTGCACTCCAAATGCCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2931
Db 361 TGTGTGCACTCCAAATGCCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420
Qy 2932 AGCATGAGCACCTGCCCCCTGAAATTCCTCAGGCTACTGTACCTGTAGTCAATGCTTGG 2991
Db 421 AGCATGAGCACCTGCCCCCTGAAATTCCTCAGGCTACTGTACCTGTAGTCAATGCTTGG 480
Qy 2992 GAGCAACACAGGCTGTGGTGTACTGTATCCAGCAATACTGGCAAGGGAATGCATA 3051
Db 481 GAGCAACACAGGCTGTGGTGTACTGTATCCAGCAATACTGGCAAGGGAATGCATA 540
Qy 3052 GAGGTTTCTATAAGACACAGTGAAGATGCTTGCAGCCCTTACAGAAATTTCTAT 3111
Db 541 GAGGTTTCTATAAGACACAGTGAAGATGCTTGCAGCCCTTACAGAAATTTCTAT 600
Qy 3112 CCACAGCCCTGCTCAATTCAGCATGTGCTAGAGACAGCAGATACAATGCTGTTTC 3171
Db 601 CCACAGCCCTGCTCAATTCAGCATGTGCTAGAGACAGCAGATACAATGCTGTTTC 660
Qy 3172 ATTCAGTGTCCAGCTTGCCA 3191
Db 661 ATTCAGTGTCCAGCTTGCCA 680

```

RESULT 5

BM783773

LOCUS

DEFINITION

BM783773

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BM783773 680 bp mRNA linear EST 05-MAR-2002  
 K-EST00611775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5',  
 mRNA sequence.

ACCESSION BM783773.1 GI:19132005

VERSION BM783773

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 680)

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS





```

RESULT 8
BM783685
LOCUS
DEFINITION K-EST0061659 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-B06 5',
mRNA sequence.
ACCESSION BM783685
VERSION BM783685.1 GI:19131917
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 673)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34 row: B column: 06
High quality sequence stop: 673.
Location/Qualifiers
FEATURES
source
1. 673
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-34-B06"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ19RP1; Site: 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dT-selected
mRNA by priming with dT-tailed vector. The dT-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation
method."
ORIGIN
Query Match 15.7%; Score 673; DB 12; Length 673;
Best Local Similarity 100.0%; Pred. No. 2.5e-310;
Matches 673; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2512 CAGCTGCGAATATGCGAGTATCTCAGAGCATGTCACAGTCCACCATGGGTC 2571
DB 1 CAGCTGCGAATATGCGAGTATCTCAGAGCATGTCACAGTCCACCATGGGTC 60
QY 2572 GGCCTTCGGAAGATCAATGCTCTCTACTGCTGGGAGATATGTCCTCCATTTACAAAT 2631
DB 61 GGCCTTCGGAAGATCAATGCTCTCTACTGCTGGGAGATATGTCCTCCATTTACAAAT 120
QY 2632 AGTTTACTACAGTGGATGCCCTCTGAGCCAGTCAGTCTGGAATTTTATCA 2691
DB 121 AGTTTACTACAGTGGATGCCCTCTGAGCCAGTCAGTCTGGAATTTTATCA 180
QY 2692 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC 2751

```

```

DB 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC 240
QY 2752 TGTGAAAGCCCTGCAAAACACACAGTCTTAAGCAGTCCCGACACCATGTCCTTGAAGGACA 2811
DB 241 TGTGAAAGCCCTGCAAAACACACAGTCTTAAGCAGTCCCGACACCATGTCCTTGAAGGACA 300
QY 2812 GCATGTGGAGATTGCACCGAGCGCAGCTCTGAGTCATGTGGTCAGCAACATGAAGCAG 2871
DB 301 GCATGTGGAGATTGCACCGAGCGCAGCTCTGAGTCATGTGGTCAGCAACATGAAGCAG 360
QY 2872 TGTGTGGACTCCAATGCTCTGAGTGGCTCCCTTCCCTTTGGCCAGTGTGGAATGGTAT 2931
DB 361 TGTGTGGACTCCAATGCTCTGAGTGGCTCCCTTCCCTTTGGCCAGTGTGGAATGGTAT 420
QY 2932 ACGATGAGCAGCCTGCCCTCCCTGAAAATTGTTTCAGGCTACTGTACCTGTAGTCTTGTG 2991
DB 421 ACGATGAGCAGCCTGCCCTCCCTGAAAATTGTTTCAGGCTACTGTACCTGTAGTCTTGTG 480
QY 2992 GAGCAACAGGCTGTGGCTGTGTTACTGATCCAGCAATATCGCAAAAGGGAATGCATA 3051
DB 481 GAGCAACAGGCTGTGGCTGTGTTACTGATCCAGCAATATCGCAAAAGGGAATGCATA 540
QY 3052 GAGGCTTCTCTATAAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTTACAGGAAATTTCTAT 3111
DB 541 GAGGCTTCTCTATAAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTTACAGGAAATTTCTAT 600
QY 3112 CCACAGCCCTGCTCAATTCAGCAGCATGTGTCTAGAGGACAGCAGATACACTGTGCTTTTC 3171
DB 601 CCACAGCCCTGCTCAATTCAGCAGCATGTGTCTAGAGGACAGCAGATACACTGTGCTTTTC 660
QY 3172 ATTCACTGTCCAG 3184
DB 661 ATTCACTGTCCAG 673
RESULT 9
BM773003
LOCUS
DEFINITION K-EST0057379 S5SNU484s1 Homo sapiens cDNA clone S5SNU484s1-3-E09
5', mRNA sequence.
ACCESSION BM773003
VERSION BM773003.1 GI:19102618
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 632)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 3 row: E column: 09
High quality sequence stop: 632.
Location/Qualifiers
FEATURES
source
1. 632
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484s1-3-E09"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"

```

```
/lab_host="Top10F"
/clone_lib="S5SNU484s1"
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dT-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation method.
After analyzing and sequencing about 2,000 - 3,000
colonies in original cDNA library, the abundant cDNAs were
selected and amplified by PCR reaction using vector region
primer including T7 promoter as 5' primer and N(dn)14 as
3' primer. The PCR products were used as template for
synthesis of biotinylated single stranded RNA by in vitro
transcription reaction. The synthesized RNA probes were
hybridized with antisense single stranded cDNAs prepared
from original library and incubated with avidin-gel.
After removing DNA-RNA hybrids by centrifuge, the
subtracted cDNA libraries were constructed by
transformation of the remaining DNA into competent cells E.
coli Top10F' with electroporation method."
```

Query Match 14.7%; Score 632; DB 12; Length 632;  
Best Local Similarity 100.0%; Pred. No. 1.1e-290;  
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2512 CAGCTCGGAATATGCAATGCTCTCAGAGCATGTCACCTTAAACCCATGGGTC 2571  
Db 1 CAGCTCGGAATATGCAATGCTCTCAGAGCATGTCACCTTAAACCCATGGGTC 60

QY 2572 GGCCTTCGGAAGATCAATGTGCTCTACTGTGTGGGAAGATATGCCCCATTACAAAT 2631  
Db 61 GGCCTTCGGAAGATCAATGTGCTCTACTGTGTGGGAAGATATGCCCCATTACAAAT 120

QY 2632 AGTTTACTACAGTGATGCGCTGTAGCCAGCATGATCTGGATTCTGTGGAAATTTATCA 2691  
Db 121 AGTTTACTACAGTGATGCGCTGTAGCCAGCATGATCTGGATTCTGTGGAAATTTATCA 180

QY 2692 GAACCCAGTACTGGGACCTGAAGGCTGCAACCTGCATCAACCCATCAATGTAGTGTC 2751  
Db 181 GAACCCAGTACTGGGACCTGAAGGCTGCAACCTGCATCAACCCATCAATGTAGTGTC 240

QY 2752 TGTGAAAGGCGCTGAAACCAACAGTCTAAGCAGTGGCGGACACCATGTGCTTGAGACA 2811  
Db 241 TGTGAAAGGCGCTGAAACCAACAGTCTAAGCAGTGGCGGACACCATGTGCTTGAGACA 300

QY 2812 GCATGTGAGATTCACACGCGGACCTCTGATGTCATGTGTGTCAGCAGCATGAGCAG 2871  
Db 301 GCATGTGAGATTCACACGCGGACCTCTGATGTCATGTGTGTCAGCAGCATGAGCAG 360

QY 2872 TGTGTGACTCCATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGATGGAATGGTAT 2931  
Db 361 TGTGTGACTCCATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGATGGAATGGTAT 420

QY 2932 ACGATGAGACCTGCCCCCTGAAATGTTCAGGCTACTGTACCTGTAGTCAATGCTTG 2991  
Db 421 ACGATGAGACCTGCCCCCTGAAATGTTCAGGCTACTGTACCTGTAGTCAATGCTTG 480

QY 2992 GAGCAACAGGCTGTGCTGTGTGCTGATGCCAGCAATCTGGCAAGGGAATGCATA 3051  
Db 481 GAGCAACAGGCTGTGCTGTGTGCTGATGCCAGCAATCTGGCAAGGGAATGCATA 540

QY 3052 GAGGGTTTCTATAAGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAATTTCTAT 3111  
Db 541 GAGGGTTTCTATAAGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAATTTCTAT 600

ORIGIN

Query Match 14.7%; Score 632; DB 12; Length 632;  
Best Local Similarity 100.0%; Pred. No. 1.1e-290;  
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3112 CCACAGCCCTCTCTCAATTCAGCATGTGTCT 3143  
Db 601 CCACAGCCCTCTCTCAATTCAGCATGTGTCT 632

RESULT 10  
LOCUS BM783264  
DEFINITION K-EST0061166 S5SNU484 Homo sapiens cDNA clone S5SNU484-32-A05 5',  
mRNA sequence.  
ACCESSION BM783264  
VERSION BM783264.1 GI:19131496  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 632)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 32 row: A column: 05  
High quality sequence stop: 632.  
Location/Qualifiers  
1..632  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S5SNU484-32-A05"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/cell\_line="SNU-484"  
/lab\_host="Top10F"  
/clone\_lib="S5SNU484"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;  
Site 2: NotI; The poly (A) + RNA was decapped with tobacco  
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
including EcoRI site by treatment of T4 RNA ligase. The  
first strand cDNA was synthesized from oligo dT-selected  
mRNA by priming with dt-tailed vector. The dt-tailed  
vector was adjusted to have about 60nt. The cDNA vector  
was circularized with E. coli DNA ligase after digestion  
of EcoRI which site is also included in vector. An RNA  
strand converted to a DNA strand by Okayama-Berg method.  
The obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F' by electroporation  
method."

ORIGIN

Query Match 14.7%; Score 632; DB 12; Length 632;  
Best Local Similarity 100.0%; Pred. No. 1.1e-290;  
Matches 632; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2512 CAGCTCGGAATATGCAATGCTCTCAGAGCATGTCACCTTAAACCCATGGGTC 2571  
Db 1 CAGCTCGGAATATGCAATGCTCTCAGAGCATGTCACCTTAAACCCATGGGTC 60

QY 2572 GGCCTTCGGAAGATCAATGTGCTCTACTGTGTGGGAAGATATGCCCCATTACAAAT 2631  
Db 61 GGCCTTCGGAAGATCAATGTGCTCTACTGTGTGGGAAGATATGCCCCATTACAAAT 120

QY 2632 AGTTTACTACAGTGATGCGCTGTAGCCAGCATGATCTGGATTCTGTGGAAATTTATCA 2691  
Db 121 AGTTTACTACAGTGATGCGCTGTAGCCAGCATGATCTGGATTCTGTGGAAATTTATCA 180

QY 2692 GAACCCAGTACTGGGACCTGAAGGCTGCAACCTGCATCAACCCATCAATGTAGTGTC 2751  
Db 181 GAACCCAGTACTGGGACCTGAAGGCTGCAACCTGCATCAACCCATCAATGTAGTGTC 240

QY 2752 TGTGAAAGGCGCTGAAACCAACAGTCTAAGCAGTGGCGGACACCATGTGCTTGAGACA 2811  
Db 241 TGTGAAAGGCGCTGAAACCAACAGTCTAAGCAGTGGCGGACACCATGTGCTTGAGACA 300

QY 2812 GCATGTGAGATTCACACGCGGACCTCTGATGTCATGTGTGTCAGCAGCATGAGCAG 2871  
Db 301 GCATGTGAGATTCACACGCGGACCTCTGATGTCATGTGTGTCAGCAGCATGAGCAG 360

QY 2872 TGTGTGACTCCATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGATGGAATGGTAT 2931  
Db 361 TGTGTGACTCCATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGATGGAATGGTAT 420

QY 2932 ACGATGAGACCTGCCCCCTGAAATGTTCAGGCTACTGTACCTGTAGTCAATGCTTG 2991  
Db 421 ACGATGAGACCTGCCCCCTGAAATGTTCAGGCTACTGTACCTGTAGTCAATGCTTG 480

QY 2992 GAGCAACAGGCTGTGCTGTGTGCTGATGCCAGCAATCTGGCAAGGGAATGCATA 3051  
Db 481 GAGCAACAGGCTGTGCTGTGTGCTGATGCCAGCAATCTGGCAAGGGAATGCATA 540

QY 3052 GAGGGTTTCTATAAGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAATTTCTAT 3111  
Db 541 GAGGGTTTCTATAAGACCAAGTGAAGATGCTTTCGCAAGCCCTTACAGGAATTTCTAT 600



```

Db      492  TTCTGTACCAAGTAAACAGTCATTTTGGACTCAGAGGAATATCTAACTTAAAGTCC 551
Qy      1019 CCAGAGCATCTCAATAAGCTGTGTGTCATGAGTAAACATTTATGTGGGTTTGTGGAGATATA 1078
Db      552  CCAGAGCATCTCAATAAGCTGTGTGTCATGAGTAAACATTTATGTGGGTTTGTGGAGATATA 611
Qy      1079 TGTTCAACCACTCAGATTATACATGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGC 1138
Db      612  TGTTCAACCACTCAGATTATACATGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGC 671
Qy      1139 TTC 1141
Db      672  TTC 674

RESULT 12
BM783788
LOCUS   K-EST0061794 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-E09 5',
DEFINITION mRNA sequence.
ACCESSION BM783788
VERSION   BM783788.1 GI:19132020
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 683)
AUTHORS  Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE    21C Frontier Korean EST Project 2001
JOURNAL  Unpublished (2002)
COMMENT  Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 36 row: E column: 09
High quality sequence stop: 683.
Location/Qualifiers
1. 683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-36-E09"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10P"
/clone_lib="S5SNU484"
/notice="Organ: Stomach; Vector: pTZ19RP1; Site: 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tabacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P by electroporation
method."

Query Match 13.5%; Score 581; DB 12; Length 683;
Best Local Similarity 99.7%; Pred. No. 2.9e-266;
Matches 681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN

```

```

Qy      2512 CAGCTGCGAATAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC 2571
Db      1    CAGCTGCGAATAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC 60
Qy      2572 GGCCTTCGGAAGATCAATGTGTCTTACTGTCTCTGGAGATATGTCCCATTTTCAAAAT 2631
Db      61  GGCCTTCGGAAGATCAATGTGTCTTACTGTCTCTGGAGATATGTCCCATTTTCAAAAT 120
Qy      2632 AGTTTACTACAGTGGATGCCGTCTGAGCCCACTGATGTGGATTCTGTGGAATTTATCA 2691
Db      121  AGTTTACTACAGTGGATGCCGTCTGAGCCCACTGATGTGGATTCTGTGGAATTTATCA 180
Qy      2692 GAACCCAGTACTCGGGACTGAAGCTGCACCTGCATCAACCCACTCAATGGTAGTGTTC 2751
Db      181  GAACCCAGTACTCGGGACTGAAGCTGCACCTGCATCAACCCACTCAATGGTAGTGTTC 240
Qy      2752 TGTGAAGGCCCTGCAAAACACACAGTGTCTAAGCAGTGTCCGACACCATGTGCTTTGAGGACA 2811
Db      241  TGTGAAGGCCCTGCAAAACACACAGTGTCTAAGCAGTGTCCGACACCATGTGCTTTGAGGACA 300
Qy      2812 GCATGTGGAGATTGCACAGCGGCGAGCTCTGAGTGCATGTGTGAGCAGACATGAAGCAG 2871
Db      301  GCATGTGGAGATTGCACAGCGGCGAGCTCTGAGTGCATGTGTGAGCAGACATGAAGCAG 360
Qy      2872 TGTGTGCACTCCAATGCCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGAAATGGTAT 2931
Db      361  TGTGTGCACTCCAATGCCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGAAATGGTAT 420
Qy      2932 AGCATGAGCACCTGCCCCCTGAAATTTGTCAGGCTACTGTACTGTAGTCAATGTGCTTGG 2991
Db      421  AGCATGAGCACCTGCCCCCTGAAATTTGTCAGGCTACTGTACTGTAGTCAATGTGCTTGG 480
Qy      2992 GAGCAACAGGCTGTGGCTGTGTACTGATCCACCAATACTGCGAAGGGGAAATGCATA 3051
Db      481  GAGCAACAGGCTGTGGCTGTGTACTGATCCACCAATACTGCGAAGGGGAAATGCATA 540
Qy      3052 GAGGGTTCTCTATAAAGGACCACTGAAGATGCCCTTCGCAAGCCCTACAGGAAATTTCTAT 3111
Db      541  GAGGGTTCTCTATAAAGGACCACTGAAGATGCCCTTCGCAAGCCCTACAGGAAATTTCTAT 600
Qy      3112 CCACAGCCCCCTGCTCAATTCAGCATGTGTCTAGAGCAGCAGATACACATCGGCTTTTC 3171
Db      601  CCACAGCCCCCTGCTCAATTCAGCATGTGTCTAGAGCAGCAGATACACATCGGCTTTTC 660
Qy      3172 ATTCACTGTCCAGCTTGCCAATG 3194
Db      661  ATTCACTGTCCAGCTTGCCAATG 683

```

```

RESULT 13
BE894518
LOCUS   60143126F1 NTH_MGC_72 Homo sapiens cDNA clone IMAGE:3918271 5',
DEFINITION mRNA sequence.
ACCESSION BE894518
VERSION   BE894518.1 GI:10356969
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 633)
AUTHORS  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE    NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL  National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT  Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

```





```

QY 2872 TGTTGGACTCCAATGCTATGTGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2931
Db 361 TGTTGGACTCCAATGCTATGTGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420

QY 2932 ACAGTAGACACTGCCCCCTGAAAATTTGTTCAAGGTACTGTACCTGTAGTCAATGCTTG 2991
Db 421 ACAGTAGACACTGCCCCCTGAAAATTTGTTCAAGGTACTGTACCTGTAGTCAATGCTTG 480

QY 2992 GAGCAACACAGCTGTGCTCGTGTACTGATCCAGCAATACTGCAAGGGAATGCATA 3051
Db 481 GAGCAACACAGCTGTGCTCGTGTACTGATCCAGCAATACTGCAAGGGAATGCATA 540

QY 3052 GAGGTTTCTTATAAGACCAAGTGAAGATCCTTGG 3087
Db 541 GAGGTTTCTTATAAGACCAAGTGAAGATCCTTGG 576

RESULT 15
BF672370 917 bp mRNA linear EST 21-DEC-2000
LOCUS 602150605F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4291680 5',
DEFINITION mRNA sequence.
ACCESSION BF672370
VERSION BF672370.1 GI:11946265
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS 1 (bases 1 to 917)
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NTH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM1136 row: j column: 01
High quality sequence stop: 563.
FEATURES
Location/Qualifiers
1..917
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4291680"
/lab_hosts="DH10B (TI phage-resistant)"
/clone_lib="NIH_MGC_81"
/notes="Organ: muscle (skeletal); Vector: pDNR-LIB
(Clontech); Site 1: SfiI (ggcgctcgcc); Site 2: SfiI
(ggccattatggc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CAGGCGCAATATGGCC-3'
and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGGCGGCACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

```

Query Match 13.0%; Score 556; DB 10; Length 917;
Best Local Similarity 100.0%; Pred. No. 2.9e-254;
Matches 556; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 2833 GGCAGCTCTGAGTGCATGTGGTGCAGCAACATCAAGCAGTGTGTGACTCCAATGCTAT 2892
Db 1 GGCAGCTCTGAGTGCATGTGGTGCAGCAACATCAAGCAGTGTGTGACTCCAATGCTAT 60
```

```

QY 2893 GTGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACTGCCCCCT 2952
Db 61 GTGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACTGCCCCCT 120

QY 2953 GAAAATTTGTTCAAGGTACTGTACCTGTAGTCAATGCTTTGAGCAACACAGGTGTGGCTGG 3012
Db 121 GAAAATTTGTTCAAGGTACTGTACCTGTAGTCAATGCTTTGAGCAACACAGGTGTGGCTGG 180

QY 3013 TGTACTGATCCAGCAATACTTGGCAAGGGAATGCATAGAGGGTTCTTATAAGGACCA 3072
Db 181 TGTACTGATCCAGCAATACTTGGCAAGGGAATGCATAGAGGGTTCTTATAAGGACCA 240

QY 3073 GTAAAGATGCCCTTCGCAAGCCCCCTACAGGAAATTTCTATCCACAGCCCCCTGCTCAATTCC 3132
Db 241 GTAAAGATGCCCTTCGCAAGCCCCCTACAGGAAATTTCTATCCACAGCCCCCTGCTCAATTCC 300

QY 3133 AGCATGTCTTAGAGCAGCAGATACAACCTGGTCTTTCTACTGTCCAGCTTGCCAA 3192
Db 301 AGCATGTCTTAGAGCAGCAGATACAACCTGGTCTTTCTACTGTCCAGCTTGCCAA 360

QY 3193 TGCAACGGCCACAGTAAATGATCAATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACC 3252
Db 361 TGCAACGGCCACAGTAAATGATCAATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACC 420

QY 3253 ACAGCAAGCATCTGCAGACCTGCATATCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGG 3312
Db 421 ACAGCAAGCATCTGCAGACCTGCATATCTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGG 480

QY 3313 AAATGTGAGCCATGCAAGTGAATGGGACGCGTCTCTGTGCAACACCAACACGCGCAAG 3372
Db 481 AAATGTGAGCCATGCAAGTGAATGGGACGCGTCTCTGTGCAACACCAACACGCGCAAG 540

QY 3373 TGCCTTCGACCAACCA 3388
Db 541 TGCCTTCGACCAACCA 556

RESULT 16
BX431593 962 bp mRNA linear EST 15-MAY-2003
LOCUS BX431593 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS0DF012YL19 5-PRIME, mRNA sequence.
ACCESSION BX431593
VERSION BX431593.1 GI:30781048
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 962)
JOURNAL Li W.B., Gruber,C., Jessee,J. and Polayes,D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequenage
BP 191 91006 EVRY cedex - France
Email: secret@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10212.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAG017B02_CS01546_1&cluster=10212.f.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAG017B02_CS01546_1.
FEATURES
Location/Qualifiers
1..962
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF012YL19"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
```



QY 2812 GCATGTGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2871  
 Db 301 GCATGTGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360  
 QY 2872 TGTGTGAGCTCAATGCGCTATGTGGCTCTCTTCCTTTTGGCCAGTGTATGGAATGGTAT 2931  
 Db 361 TGTGTGAGCTCAATGCGCTATGTGGCTCTCTTCCTTTTGGCCAGTGTATGGAATGGTAT 420  
 QY 2932 ACGATGAGCACTGCCGCCCTGAAATTTGTTTCAGGCTACTGTACTCTGATGCTTGGCTTG 2991  
 Db 421 ACGATGAGCACTGCCGCCCTGAAATTTGTTTCAGGCTACTGTACTCTGATGCTTGGCTTG 480  
 QY 2992 GAGCA 2996  
 Db 481 GAGCA 485

RESULT 18  
 BI850289  
 LOCUS  
 DEFINITION imagec\_11 2000/sly246bdr81.y1 NIH\_MGC 56 Homo sapiens cDNA clone  
 IMAGE:4284387 5', mRNA sequence.  
 ACCESSION BI850289  
 VERSION BI850289  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 591)  
 Kalle, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and  
 Prange, C.K.  
 TITLE The I.M.A.G.E. Consortium quality control effort: clone  
 resequencing for verification  
 JOURNAL Unpublished (2001)  
 COMMENT Other ESTs: BF700035  
 Contact: Prange CK  
 The I.M.A.G.E. Consortium  
 Lawrence Livermore National Laboratory  
 Livermore, CA, USA  
 Email: help@image.llnl.gov

This read has been verified (found to hit its original self in the  
 correct orientation), as part of the I.M.A.G.E. Consortium quality  
 control effort. High quality sequence is defined as having 100 or  
 more base pairs with a phred quality value of 20 or greater, where  
 a sliding window of 4 base pairs marks the beginning and end of the  
 sequence. For information on obtaining this clone, please contact  
 info@image.llnl.gov.  
 Plate: L1CM117 row: j column: 4  
 Seq primer: -21n13  
 High quality sequence stop: 591.  
 Location/Qualifiers

## FEATURES

source

```

1. .591
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4284387"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC 56"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:  

SfiI (ggccgctcgcc); Site_2: SfiI (ggccattatggcc);  

Double-stranded cDNA was prepared from cell line RNA. 5'  

and 3' adaptors were used in cloning as follows: 5'  

adaptor sequence: 5'-CAGGCGCATATGGC-3' and 3' adaptor  

sequence: 5'-ATTCTAGGCGCGGCGGCATG-dT(30)BN-3'  

(where B = A, C, or G and N = A, C, G, or T). Average  

insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies  

contained inserts by PCR. This library was enriched for  

full-length clones and was constructed by Clontech  

Laboratories (Palo Alto, CA)."
```

## ORIGIN

Query Match 11.0%; Score 471; DB 12; Length 591;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-213;  
 Matches 591; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 QY 2636 TACTACAGTGGATGCGCTCTGAGCCGAGTGATGCTGGATTTCTGTGGAATTTATCAGAAC 2695  
 Db 1 TACTACAGTGGATGCGCTCTGAGCCGAGTGATGCTGGATTTCTGTGGAATTTATCAGAAC 60  
 QY 2696 CAGTACTCGGGGACTGGAAGGCTGGAACCTGCATCAACCCACTCAATGCTAGTGTCTGTG 2755  
 Db 61 CAGTACTCGGGGACTGGAAGGCTGGAACCTGCATCAACCCACTCAATGCTAGTGTCTGTG 120  
 QY 2756 AAAGGCTTCAAAACCAACAGTGTCTAAAGCAGTCCCGGACACCATGTGCTTTGAGGACAGCAT 2815  
 Db 121 AAAGGCTTCAAAACCAACAGTGTCTAAAGCAGTCCCGGACACCATGTGCTTTGAGGACAGCAT 179  
 QY 2816 GTGGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2875  
 Db 180 GTGGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 239  
 QY 2876 TGGACTTCCCAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATTTATCAGAAC 2935  
 Db 240 TGGACTTCCCAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATTTATCAGAAC 299  
 QY 2936 TGAGCACTGCCCCCTGAAATTTGTTTACGGCTACTGTACCTGTAGTCTATGCTTTGGAGC 2995  
 Db 300 TGAGCACTGCCCCCTGAAATTTGTTTACGGCTACTGTACCTGTAGTCTATGCTTTGGAGC 359  
 QY 2996 AACCAAGGCTGTGGTGTGTACTGTATCCAGCAATCTGGCAAGGGAATGCATAGAGG 3055  
 Db 360 AACCAAGGCTGTGGTGTGTACTGTATCCAGCAATCTGGCAAGGGAATGCATAGAGG 419  
 QY 3056 GTTCTTATAAAGGACCAAGTGAAGATGCCTTCCCAAGCCCTACAGGAAATTTCTATCCAC 3115  
 Db 420 GTTCTTATAAAGGACCAAGTGAAGATGCCTTCCCAAGCCCTACAGGAAATTTCTATCCAC 479  
 QY 3116 AGCCCTGTCTCAATTCACGATGTCTGTAGAGGACAGAGATACACTGGTCTTTTCATTC 3175  
 Db 480 AGCCCTGTCTCAATTCACGATGTCTGTAGAGGACAGAGATACACTGGTCTTTTCATTC 539  
 QY 3176 ACTGTCCAGCTTGGCAATGCAAGCGGCACAGTAAATGTCATCAATCAGAGCAT 3227  
 Db 540 ACTGTCCAGCTTGGCAATGCAAGCGGCACAGTAAATGTCATCAATCAGAGCAT 591

## RESULT 19

BF853006

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

BF853006 598 bp mRNA linear EST 16-JAN-2001  
 MR2-EN0093-191200-001-b07 EN0093 Homo sapiens cDNA, mRNA sequence.  
 BF853006  
 BF853006.1 GI:12240854  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 598)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 20202663  
 10737800  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?l1=MR2&t2=MR2-EN0093-191200-001-b07&t3=2000-12-19&t4=1>)  
Seq primer: puc 18 forward  
High quality sequence stop: 598.

Location/Qualifiers  
1..598

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="EN0093"  
/note="Organ: lung normal; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

ORIGIN

Query Match 10.1%; Score 435; DB 10; Length 598;  
Best Local Similarity 99.8%; Pred. No. 2.3e-196;  
Matches 485; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3282 TGGCTTCTACGGTATCCCAATGAGGGAATGTCAGCCATGCAAGTGCATATGGCA 3341  
Db 5 TGGCTTCTACGGTATCCCAATGAGGGAATGTCAGCCATGCAAGTGCATATGGCA 64

QY 3342 CGCGTCTCTGACACACACACGCGGCAAGTGTCTTGCACACACACGCGGCTCAAGGG 3401  
Db 65 CGCGTCTCTGACACACACACGCGGCAAGTGTCTTGCACACACACGCGGCTCAAGGG 124

QY 3402 GGACGAGTCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAACCCCTCTCAGAGGAAC 3461  
Db 125 GGACGAGTCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAACCCCTCTCAGAGGAAC 184

QY 3462 ATGTATTATATCTCTTATTGACTATCAGTTCACTTGTAGTCTATCCAGGAGATGA 3521  
Db 185 ATGTATTATATCTCTTATTGACTATCAGTTCACTTGTAGTCTATCCAGGAGATGA 244

QY 3522 TCGCTATTACACAGCTATCAATTTTGGCTACTCTGACCAACCAACAGGGATTGGA 3581  
Db 245 TCGCTATTACACAGCTATCAATTTTGGCTACTCTGACCAACCAACAGGGATTGGA 304

QY 3582 CATGTTATCAATGCTCCCAAGATTTCAACTCAACATCCTGGGCTGCCAGTTTCTC 3641  
Db 305 CATGTTATCAATGCTCCCAAGATTTCAACTCAACATCCTGGGCTGCCAGTTTCTC 364

QY 3642 AGCTGGACCCAGCTGAGAGAGATGCTGTTGTTTCAAAACCAACATTAGGAGTA 3701  
Db 365 AGCTGGACCCAGCTGAGAGAGATGCTGTTGTTTCAAAACCAACATTAGGAGTA 424

QY 3702 CAAGATAGTTTCTCTAATAGAGATTGATTTTTCGAACCAACCAATATCACTTTCTT 3761  
Db 425 CAAGATAGTTTCTCTAATAGAGATTGATTTTTCGAACCAACCAATATCACTTTCTT 484

QY 3762 TGTGTA 3767  
Db 485 TGTGTA 490

RESULT 20  
CB242644/C  
LOCUS  
DEFINITION  
UI-CF-FNO-afs-i-02-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone  
CB242644  
ACCESSION  
CB242644

CB242644.1 GI:28364288

EST.  
Homo sapiens (human)

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: McCray, PB

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

CDNA Library preparation: Dr. M. Bento Soares, University of Iowa

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com) or from Open Biosystems

(www.openbiosystems.com).

The following repetitive elements were found in this cDNA

sequence: 16-237, >L1P2#LINE/L1 (matched complement)

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

1..746

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-CF-FNO-afs-i-02-0-UI"

/tissue\_type="Human Lung Epithelial cells"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone\_lib="UI-CF-FNO"

/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-CF-FNO is a subtracted cDNA library derived from two

normalized Human lung epithelial cell libraries (EN1 and

DUI) The library was subtracted according to according to

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. For additional information, contact:

bento-soares@uiowa.edu

TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS

6hr to LPS 24h

TAG\_LIB=UI-CF-FNO

TAG\_SEQ=CTGCTCAGGT

ORIGIN

Query Match 9.9%; Score 426; DB 14; Length 746;

Best Local Similarity 100.0%; Pred. No. 4.9e-192;

Matches 426; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3376 TTCTGCACCAACGAGGGCTCAAGGGGACGAGTCCAGCTATGAGGTAGAAATCGA 3435

Db 661 TTCTGCACCAACGAGGGCTCAAGGGGACGAGTCCAGCTATGAGGTAGAAATCGA 602

QY 3436 TACCAAGGAACCCCTCTCAGAGGAACATGTTATTACTCTTCTTATTGACTATCAGTTTC 3495

Db 601 TACCAAGGAACCCCTCTCAGAGGAACATGTTATTACTCTTCTTATTGACTATCAGTTTC 542

QY 3496 ACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACT 3555

Db 541 ACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACT 482

QY 3556 CCTGACGACAAACAGGAGTTTGACATGTTTCATCAATGCTCCCAAGATTTCAACCTC 3615

Db 481 CCTGACGACAAACAGGAGTTTGACATGTTTCATCAATGCTCCCAAGATTTCAACCTC 422

```
QY 3616 AACATCACTGGGTGCCAGTTTCTCAGCTGGAACCCAGCTGGAGAAGAGATGCCTGTT 3675
|
|
|
Db 421 AACATCACTGGGTGCCAGTTTCTCAGCTGGAACCCAGCTGGAGAAGAGATGCCTGTT 362
|
|
|
QY 3676 GTTTCAAAACCAACATAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTTGATTTT 3735
|
|
|
Db 361 GTTTCAAAACCAACATAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTTGATTTT 302
|
|
|
QY 3736 CGCAACCAACCAATATACATCTTTCTTTTATGTAGTCAATTTCACTGCGCCCATCAA 3795
|
|
|
Db 301 CGCAACCAACCAATATACATCTTTCTTTTATGTAGTCAATTTCACTGCGCCCATCAA 242
|
|
|
QY 3796 ATTACAG 3801
|
|
|
Db 241 ATTACAG 236
|
|
|
RESULT 21
BF853681
LOCUS BF853681 595 bp mRNA linear EST 16-JAN-2001
DEFINITION MR2-EN0093-211200-003-a05 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853681
VERSION BF853681.1 GI:12241425
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 595)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.P., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.K., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
211200-003-a05&t3=2000-12-21&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 595.
FEATURES
Location/Qualifiers
1..595
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="adult"
/clone_lib="EN0093"
/note="Organ: lung normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
ORIGIN
Query Match 9.9%; Score 425; DB 10; Length 595;
Best Local Similarity 99.8%; Pred. No. 1.4e-151;
```

```
Matches 475; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 3282 TGGCTTTACGGTGTATCCACCAATGGAGGAAATGTTCAGCCATCAATGTGCAATGGGCA 3341
|
|
|
Db 6 TGGCTTTACGGTGTATCCACCAATGGAGGAAATGTTCAGCCATCAATGTGCAATGGGCA 65
|
|
|
QY 3342 CCGCTCTCTGTGCAACACCAACACCGGCAAGTGTCTTCGACCACCAAGGGGCTCAAGG 3401
|
|
|
Db 66 CCGCTCTCTGTGCAACACCAACACCGGCAAGTGTCTTCGACCACCAAGGGGCTCAAGG 125
|
|
|
QY 3402 CGACGAGTGCACAGCTATGTGAGGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAAC 3461
|
|
|
Db 126 GACGAGTGCACAGCTATGTGAGGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAAC 185
|
|
|
QY 3462 ATGTTATTATATCTCTCTTATTGACTATCAGTTACCTTTAGTCTATCCAGGAAGATGA 3521
|
|
|
Db 186 ATGTTATTATATCTCTCTTATTGACTATCAGTTACCTTTAGTCTATCCAGGAAGATGA 245
|
|
|
QY 3522 TCGCTATTACACAGCTATCAATTTTGTGGTACTCTCTGACGAACAAACACAGGATTTGA 3581
|
|
|
Db 246 TCGCTATTACACAGCTATCAATTTTGTGGTACTCTCTGACGAACAAACACAGGATTTGA 305
|
|
|
QY 3582 CATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCAGTTTCTC 3641
|
|
|
Db 306 CATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCAGTTTCTC 365
|
|
|
QY 3642 AGCTGGACCCAGGCTGGAGAGAGATGCTGTTTCTTCAAAACACCAATTAAGGATGA 3701
|
|
|
Db 366 AGCTGGACCCAGGCTGGAGAGAGATGCTGTTTCTTCAAAACACCAATTAAGGATGA 425
|
|
|
QY 3702 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCAACCAATATACATT 3757
|
|
|
Db 426 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCAACCAATATACATT 481
|
|
|
RESULT 22
BF183626/c
LOCUS BF183626 744 bp mRNA linear EST 15-JUL-2003
DEFINITION UI-R-EU0-azn-a-22-0-UI-s1 NCI-CGAP_Carl Homo sapiens cDNA clone
IMAGE:5850909 3', mRNA sequence.
ACCESSION BF183626
VERSION BF183626.1 GI:20359183
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 744)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
The following repetitive elements were found in this cDNA
sequence: 16-237, >hIPAZ#LINE/L1 (matched complement)
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES
Location/Qualifiers
1..744
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5850909"
/tissue_type="Osteoarthritic Cartilage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
```



```

QY 3766 TATGTCAGTAATTTCACTGGCCC 3789
Db 87 TATGTCAGTAATTTCACTGGCCC 64

RESULT 24
BF853962 BR853962 544 bp mRNA linear EST 16-JAN-2001
LOCUS MR2-EN0093-261200-004-g08 EN0093 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF853962
ACCESSION BF853962.1 GI:12241706
VERSION BF853962.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 544)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
261200-004-g08&t3=2000-12-26&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 544.
Location/Qualifiers
1..544
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0093"
/note="Organ: lung normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

FEATURES
source
Query Match 9.3%; Score 400; DB 10; Length 544;
Best Local Similarity 99.8%; Pred. No. 1.3e-179;
Matches 450; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3336 TGGGACGCGTCTGTGACACACACAGGGGAGGTGTTTGGACACCAAGGGCGT 3395
Db 63 TGGGACGCGTCTGTGACACACACAGGGGAGGTGTTTGGACACCAAGGGCGT 122

QY 3396 CAAGGGGACAGTGCAGCTATGTAGGTAGAAAATCGATACCAAGGAACCTCTTCAG 3455
Db 123 CAAGGGGACAGTGCAGCTATGTAGGTAGAAAATCGATACCAAGGAACCTCTTCAG 182

QY 3456 AGGAACATGTATTATCTCTTATTGACTATCAGTTCACCTTTAGTCTATCCACGGA 3515
Db 183 AGGAACATGTATTATCTCTTATTGACTATCAGTTCACCTTTAGTCTATCCACGGA 242

```

```

QY 3516 AGATGATCGCTATTACACAGCTATCAATTTTGGTACTCTCTGACGACAAACAGGGA 3575
Db 243 AGATGATCGCTATTACACAGCTATCAATTTTGGTACTCTCTGACGACAAACAGGGA 302

QY 3576 TTTCGACATGTTTCATCAATGCTCCCAAGAAATTTCAACCTCAACATCACCCTGGCTGCCAG 3635
Db 303 TTTCGACATGTTTCATCAATGCTCCCAAGAAATTTCAACCTCAACATCACCCTGGCTGCCAG 362

QY 3636 TTTCCTCAGCTGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTA 3695
Db 363 TTTCCTCAGCTGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTA 422

QY 3696 GGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTCATTTTCGCAACCCCAAAATATCAC 3755
Db 423 GGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTCATTTTCGCAACCCCAAAATATCAC 482

QY 3756 TTTCCTTTGTTTATGTCAGTAATTTCACTGG 3786
Db 483 TTTCCTTTGTTTATGTCAGTAATTTCACTGG 513

RESULT 25
CA306924/c
LOCUS CA306924
DEFINITION UI-H-Ftl1-bhu-e-21-0-UI-s1 NCI CGAP Ftl1 Homo sapiens cDNA clone
UI-H-Ftl1-bhu-e-21-0-UI 3', mRNA sequence.
ACCESSION CA306924
VERSION CA306924.1 GI:24469978
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 643)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Gary W. Hunninghake, U of I
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bentso-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 16-237 >L1PA2#LINE/L1 (matched complement)
Seq primer: M13 FORWARD
POLYA=yes.
Location/Qualifiers
1..643
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="NCI CGAP Ftl1"
/tissue_type="Aveolar Macrophage"
/lab_host="DH10B (Life Technologies)"
/dev_stage="Adult"
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI-CGAP Ftl1 is a normalized cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The library was
normalized according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. First strand cDNA synthesis was
primed with an oligo-dT primer containing a Not I site.
Double stranded cDNA was ligated to an EcoR I adaptor,
digested with Not I, and cloned directionally into
pT73-Pac vector. The oligonucleotide used to prime the
synthesis of first-strand cDNA contains a library tag

```

sequence that is located between the Not I site and the  
GTCATC tail. The sequence tag for this library is  
GGCCATGCCG. The tissue was provided by Dr. Gary W.  
Hunninghake of the University of Iowa.  
TAG TISSUE=Human Lung Atelect Macrophage  
TAG\_LIB=UI-H-FT1  
TAG\_SEQ=GGCCATGCCG"

## ORIGIN

Query Match	9.3%;	Score 399;	DB 14;	Length 643;
Best Local Similarity	100.0%;	Pred. No. 4e-179;		
Matches 399;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
3403	GACGAGTCCGAGCTATGTGAGGTAGAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAACA	3462		
DB				
634	GACGAGTCCGAGCTATGTGAGGTAGAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAACA	575		
3463	TGTTATTATACCTCTCTTATTAGTACTATCAGTTTCACCTTTAGTCTATCCGAGGAAGATGAT	3522		
DB				
574	TGTTATTATACCTCTCTTATTAGTACTATCAGTTTCACCTTTAGTCTATCCGAGGAAGATGAT	515		
3523	CGCTATTACAGAGTATCAATTTTGTGGCTACTCCTGACGACACAAACAGGGATTTGCGAC	3582		
DB				
514	CGCTATTACAGAGTATCAATTTTGTGGCTACTCCTGACGACACAAACAGGGATTTGCGAC	455		
3583	ATGTTTCATCAATGCTCCAGAAATTTCAAGCTCAACATCACCTGGGCTGCCAGTTTCTCA	3642		
DB				
454	ATGTTTCATCAATGCTCCAGAAATTTCAAGCTCAACATCACCTGGGCTGCCAGTTTCTCA	395		
3643	GCTGGAACCCGAGGTGAGAGAGATGCTGTGTGTTTCAAAAACCAACATTAAAGAGTAC	3702		
DB				
394	GCTGGAACCCGAGGTGAGAGAGATGCTGTGTGTTTCAAAAACCAACATTAAAGAGTAC	335		
3703	AAAGATAGTTTCTCTAATAGAGAAGTTTGATTTTCGCAACACCCCAAAATATCACTTTCTTTT	3762		
DB				
334	AAAGATAGTTTCTCTAATAGAGAAGTTTGATTTTCGCAACACCCCAAAATATCACTTTCTTTT	275		
3763	GTTTATGTCAGTAATTTTCACCTGGCCCATCAAAATTCAG	3801		
DB				
274	GTTTATGTCAGTAATTTTCACCTGGCCCATCAAAATTCAG	236		

RESULT 26	AY418588	3671 bp	DNA	linear	GSS 17-DEC-2003
LOCUS	Pan troglodytes ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,				
DEFINITION	genomic survey sequence.				
ACCESSION	AY418588				
VERSION	AY418588.1	GI:39774548			
KEYWORDS	GSS.				
SOURCE	Pan troglodytes (chimpanzee)				
ORGANISM	Pan troglodytes				
REFERENCE	1 (bases 1 to 3671)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5552), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 3671)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	These sequences were made by sequencing genomic exons and ordering them based on alignment.				



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 739)  
Bonald,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
8889548  
Contact: McCray, PB  
McCray Lab  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: paul-mccray@iowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).  
The following repetitive elements were found in this cDNA  
sequence: 16-237, >LIPA2#LINE/L1 (matched complement)  
Seq primer: M13 FORWARD  
POLYA=Yes.

FEATURES  
source  
1..739  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-CF-FNO-afg-d-10-0-UI"  
/tissue\_type="Human Lung Epithelial cells"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-CF-FNO"  
/note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; UI-CF-FNO is a subtracted cDNA library derived from two normalized Human lung epithelial cell libraries (EN1 and DUI). The library was subtracted according to according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. For additional information, contact: benton-soares@uiowa.edu  
TAG\_TISSUE=Human Lung Epithelial Cell Lines untreated LPS 6hr to LPS 24h  
TAG\_L1B=UI-CF-FNO  
TAG\_SEQ=CTGCTCAGGT"

ORIGIN  
Query Match 9.1%; Score 389; DB 14; Length 739;  
Best Local Similarity 100.0%; Pred. No. 2.5e-174;  
Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3413 AGCTATGCTGAGGTAGAAAATCGATACCAAGAAACCTCTCAGAGGAACATGTTATTATA 3472  
DB 624 AGCTATGCTGAGGTAGAAAATCGATACCAAGAAACCTCTCAGAGGAACATGTTATTATA 3472  
QY 3473 CTCTTCTTATGACTATACGTTACCTTTAGCTTTATGCCAGGAAGATGCTGTTATACA 3532  
DB 564 CTCTTCTTATGACTATACGTTACCTTTAGCTTTATGCCAGGAAGATGCTGCTATTACA 505  
QY 3533 CAGCTATCAATTTTGTGGCTACTCTGACGAAACAAACAGGATTTGGACATGTTCTATCA 3592  
DB 504 CAGCTATCAATTTTGTGGCTACTCTGACGAAACAAACAGGATTTGGACATGTTCTATCA 445  
QY 3593 ATGCCTCCAGAAATTTCAACCTCAACATCACTCTGGGCTGCCAGTTTCTCAGCTGGAACCC 3652  
DB 444 ATGCCTCCAGAAATTTCAACCTCAACATCACTCTGGGCTGCCAGTTTCTCAGCTGGAACCC 385  
QY 3653 AGGCTGAGAGAGATGCTGTTTCAAAACCAACATTAGGAGTACAAAGATAGTT 3712  
DB 384 AGGCTGAGAGAGATGCTGTTTCAAAACCAACATTAGGAGTACAAAGATAGTT 325

QY 3713 TCTCTAATGAGAAGTTTGTGATTTTTCGCAACCAACCAAAATATCACTTTCTTTTGTATGCA 3772  
DB 324 TCTCTAATGAGAAGTTTGTGATTTTTCGCAACCAACCAAAATATCACTTTCTTTTGTATGCA 265  
QY 3773 GTAAATTTCACTGGCCCATCAAAATTCAG 3801  
DB 264 GTAAATTTCACTGGCCCATCAAAATTCAG 236

RESULT 28  
BG563729 674 bp mRNA linear EST 10-APR-2001  
LOCUS 602584512F1 NIH\_MGC\_76 Homo sapiens cDNA clone IMAGE:4712114 5',  
mRNA sequence.  
ACCESSION BG563729 GI:13571381  
VERSION BG563729.1  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 674)  
NIH-MGC http://mgi.nci.nih.gov/  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: CLONETECH Laboratories, Inc.  
CDNA Library Preparation: CLONETECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLCW1553 row: h column: 03  
High quality sequence stop: 674.  
Location/Qualifiers  
1..674  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4712114"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NIH\_MGC\_76"  
/note="Organ: Liver; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccattcgcc); Site 2: SfiI (ggccattcgcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATATGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGCGCGCATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.85 kb (range 1.0-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 8.9%; Score 382; DB 12; Length 674;  
Best Local Similarity 99.6%; Pred. No. 5.6e-171;  
Matches 552; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 2568 GGTGCGCCTTCGGAAGATCAATGTCTCTACTGTGTGGGAAGATATGCCCATTTTAC 2627  
DB 1 GGTGCGCCTTCGGAAGATCAATGTCTCTACTGTGTGGGAAGATATGTCCCCATTTC 60  
QY 2628 AAATAGTTTACTACAGTGGATGCGCTGAGCCGAGTATGCTGTGGAAATTTT 2687  
DB 61 AAATAGTTTACTACAGTGGATGCGCTGAGCCGAGTATGCTGTGGAAATTTT 120  
QY 2688 ATCAGAACCCAGTACTCTGGGAGCTCAAGCTGCACCTGCATCAACCCACATCAATGGTAG 2747  
DB 121 ATCAGAACCCAGTACTCTGGGAGCTCAAGCTGCACCTGCATCAACCCACATCAATGGTAG 180

```

QY 2748 TGTCTGTGAAGGCTGCAAAACACACAGTGTCTAAGCAGTGCAGCAGCAGTGTGCCTTGAG 2807
Db 181 TGTCTGTGAAGGCTGCAAAACACACAGTGTCTAAGCAGTGCAGCAGCAGTGTGCCTTGAG 240
QY 2808 GACAGCATGTGGAGATTGCCACAGCGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAA 2867
Db 241 GACAGCATGTGGAGATTGCCACAGCGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAA 300
QY 2868 GCAGTGTGTGACTCCAAATCCTATGCTGCTCTCTCCCTTT-TGGCCAGTGTATGGAAT 2926
Db 301 GCAGTGTGTGACTCCAAATCCTATGCTGCTCTCTCCCTTTGTCGCGAGTGTATGGAAT 360
QY 2927 GGTATACGATGAGCAGCTGCCGCCCTGAAATTTGTTGAGGCTACTGTACCTGTGAGTCATT 2986
Db 361 GGTATACGATGAGCAGCTGCCGCCCTGAAATTTGTTGAGGCTACTGTACCTGTGAGTCATT 420
QY 2987 GCTTGGAGCAACAGGCTGTGGTGTGTTACTGATCCAGCAATATCTGGCAAGGGAAT 3046
Db 421 GCTTGGAGCAACAGGCTGTGGTGTGTTACTGATCCAGCAATATCTGGCAAGGGAAT 480
QY 3047 GCATAGAGGTTCTTATAAAGGACCACTGAAGATGCTTCGCAAGCCCTTACAGGAAAT 3106
Db 481 GCATAGAGGTTCTTATAAAGGACCACTGAAGATGCTTCGCAAGCCCTTACAGGAAAT 540
QY 3107 TCTATCCACAGCCC 3120
Db 541 TCTATCCACAGCCC 554

RESULT 29
LOCUS AL048842 502 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFP43462118_r1_434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL048842
VERSION DKFP43462118, mRNA sequence.
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 502)
AUTHORS Ootenwaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
TITLE EST (Ootenwaelder, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by MediGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No si sequence
available.
This clone (DKFP43462118) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..502
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP43462118"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
Query Match 8.6%; Score 369; DB 9; Length 502;
Best Local Similarity 99.6%; Pred. No. 8.8e-165;

```

```

Matches 469; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 474 CACGTGGCTCATTCGAGGACAGCCAAATAGAAATAGAGCTTCGTTCAATCATTTTGC 533
Db 1 CACGTGGCTCATTCGAGGACAGCCAAATAGAAATAGAGCTTCGTTCAATCATTTTGC 60
QY 534 TACAGAGTGTAGTTGGGACCACTTTATATGTTTATGATGGGACTCAATTTATGACCCGCT 593
Db 61 TACAGAGTGTAGTTGGGACCACTTTATATGTTTATGATGGGACTCAATTTATGACCCGCT 120
QY 594 AGTTGCTCATTTAGTGGCCCTCATGTTCTTCGAGAGATGGCAATGAGAGTGTCCCTGA 653
Db 121 AGTTGCTCATTTAGTGGCCCTCATGTTCTTCGAGAGATGGCAATGAGAGTGTCCCTGA 180
QY 654 GGTTCGTCACATCAGGTTATGCTTGTGCTCAATTTTGTAGTGTGCTTATAATTT 713
Db 181 GGTTCGTCACATCAGGTTATGCTTGTGCTCAATTTTGTAGTGTGCTTATAATTT 240
QY 714 GACTGGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 773
Db 241 GACTGGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 300
QY 774 AGAGTGTGAAGATCAGTAATAGCAGCAAACTGTTGAATGTGAATGTCTGAAACTGGAA 833
Db 301 AGAGTGTGAAGATCAGTAATAGCAGCAAACTGTTGAATGTGAATGTCTGAAACTGGAA 360
QY 834 AGGTGAAGCATGTGACATTTCTCTACTGTACAGCAAACTGTTGTTTCTCATCGAGCAT 893
Db 361 AGGTGAAGCATGTGACATTTCTCTACTGTACAGCAAACTGTTGTTTCTCATCGAGCAT 420
QY 894 CTGCAATTCAGTGTGATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAGG 944
Db 421 CTGCAATTCAGTGTGATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAGG 471

RESULT 30
LOCUS BE880905 991 bp mRNA linear EST 20-OCT-2000
DEFINITION BE880905_1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3892859 5',
mRNA sequence.
ACCESSION BE880905
VERSION BE880905.1 GI:10329681
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 991)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/DRP/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAN9679 row: p column: 12
High quality sequence stop: 607.
FEATURES
source
1..991
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="large cell carcinoma, undifferentiated"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_69"
/notes="Organ: lung; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.

```

```
ORIGIN
Average insert size 1.1 kb. Library constructed by Life
Technologies."
Query Match 8.5%; Score 363; DB 10; Length 991;
Best Local Similarity 99.6%; Pred. No. 7.6e-162; Indels 1; Gaps 1;
Matches 533; Conservative 0; Mismatches 1;
QY 1536 GGCCCTATACGTTTCATGGTGGCTTACAAAGGCTTTTCAGTGGCAATPAAGTACCGGCTTCGAGA 1595
DB 38 GGCCCTATACGTTTCATGGTGGCTTACAAAGGCTTTTCAGTGGCAATPAAGTACCGGCTTCGAGA 97
QY 1596 TGATCTCTACGATATAGTGTGATACCCAGATGTCGACCATCTTAAAGACAGCGGAT 1655
DB 98 TGATCTCTACGATATAGTGTGATACCCAGATGTCGACCATCTTAAAGACAGCGGAT 157
QY 1656 TTTCCGTTACTTGCACACAGCTGTGATAGTGAAGTGAACCATCTCTGTGTTTGGGGGAAA 1715
DB 158 TTTCCGTTACTTGCACACAGCTGTGATAGTGAAGTGAACCATCTCTGTGTTTGGAGAAA 217
QY 1716 CACACAAATGACATCTATAGGCAATGGCGCAATGTTCTCTTCAGATTTATGGC 1775
DB 218 CACACAAATGACATCTATAGGCAATGGCGCAATGTTCTCTTCAGATTTATGGC 277
QY 1776 CTATGACATTCGCTGTGACCGCTGTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGT 1835
DB 278 CTATGACATTCGCTGTGACCGCTGTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGT 337
QY 1836 CAACAGATTTGGCCATTCAGACGCTTTACACACAGCACCATGATGTGTGTTGGTGGTTT 1895
DB 338 CAACAGATTTGGCCATTCAGACGCTTTACACACAGCACCATGATGTGTGTTGGTGGTTT 397
QY 1896 CAATAGTCTCCTCTCAGACGACATCTGGTATTCACCTCGGACAGCTGTG-ATGCCATC 1954
DB 398 CAATAGTCTCCTCTCAGACGACATCTGGTATTCACCTCGGACAGCTGTGATTCGCGATC 457
QY 1955 GGAGTGAAGCGCTTGTGTAGCAGGACCTGATTCGGTGTGTGTGGAACACAGGTT 2014
DB 458 GGAGTGAAGCGCTTGTGTAGCAGGACCTGATTCGGTGTGTGTGGAACACAGGTT 517
QY 2015 CGTCTCAGTATCTCGTGGCGCTGCAACTGATGACAGAAAGTTTAAA 2069
DB 518 CGTCTCAGTATCTCGTGGCGCTGCAACTGATGACAGAAAGTTTAAA 572
RESULT 31
BF963698/c 484 bp mRNA linear EST 22-JAN-2001
LOCUS PM4-NN1204-221200-001-g03 NN1204 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF963698.1 GI:12380973
VERSION BF963698.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Nagai M.A., da Silva W. Jr., Zago M.A., Bordin S., Costa F.,
Dias Neto E., Garcia Correa R., Verjovski-Almeida S., Briones M.R.,
Goldman G.H., Carvalho A.P., Matsukuma A., Baia G.S., Simpson D.H.,
Brunstein A., Gellivelira P.S., Bucher F., Jongeneel C.V.,
O'Hare M.J., Soares F., Brentani R.R., Reis L.F., de Souza S.J. and
Simpson A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
```

Tel: +55-11-2704922  
Fax: +55-11-2707001

Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM4&t2=PM4-NN1204-221200-001-g03&t3=2000-12-22&t4=1>)

Seq primer: puc 18 forward  
High quality sequence start: 8  
High quality sequence stop: 350.

Location/Qualifiers  
1. 484

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="NN1204"

/note="Organ: nervous normal; Vector: puc18; Site: 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORSTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

#### ORIGIN

Query Match 8.4%; Score 360; DB 10; Length 484;  
Best Local Similarity 99.8%; Pred. No. 1.8e-160;  
Matches 410; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3071 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATT 3130  
DB 484 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATT 425  
QY 3131 CCAGCATGTGCTAGAGCAGCAGATACAACTGGTCTTTCATTCACATGTCAGCTGCC 3190  
DB 424 CCAGCATGTGCTAGAGCAGCAGATACAACTGGTCTTTCATTCACATGTCAGCTGCC 365  
QY 3191 AATGCAAGCGGCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGA 3250  
DB 364 AATGCAAGCGGCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGA 305  
QY 3251 CCACAGGCAAGCAGTGCAGAGCTGCATATCTGGCTTCTACGGTATCCACCATATGAG 3310  
DB 304 CCACAGGCAAGCAGTGCAGAGCTGCATATCTGGCTTCTACGGTATCCACCATATGAG 245  
QY 3311 GGAATGTGAGCATGCAAGTGAATGGCAGCGCTCTGTGCAACACCAACACCGGCA 3370  
DB 244 GGAATGTGAGCATGCAAGTGAATGGCAGCGCTCTGTGCAACACCAACACCGGCA 185  
QY 3371 ACTGCTTCTGCAACCAAGGGCGTCAAGGGGGAGCGAGTGCAGCTATGTAGGTAGAAA 3430  
DB 184 ACTGCTTCTGCAACCAAGGGCGTCAAGGGGGAGCGAGTGCAGCTATGTAGGTAGAAA 125  
QY 3431 ATCGATACCAAGGAAACCTCTTCAGAGGAACATGTTATATCTTCTTA 3481  
DB 124 ATCGATACCAAGGAAACCTCTTCAGAGGAACATGTTATATCTTCTTA 74

BF1313142 953 bp mRNA linear EST 21-NOV-2000  
601897661F1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:4126959 5',  
mRNA sequence.

ACCESSION BF1313142

VERSION BF1313142.1 GI:11261090

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 953)

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.

**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: ATCC  
 cDNA Library Preparation: Ling Hong/Rubin Laboratory  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
 Plate: LLM1017 row: j column: 16  
 High quality sequence stop: 612.  
**FEATURES** Location/Qualifiers  
 source  
 1..953  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4126959"  
 /tissue\_type="neuroblastoma"  
 /lab\_host="DH10B (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_19"  
 /note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."  
**ORIGIN**  
 Query Match 8.3%; Score 355; DB 10; Length 953;  
 Best Local Similarity 99.6%; Pred. No. 5.1e-158;  
 Matches 525; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
 QY 1180 GGTCAATCTTTGGCATTATACAGGATAAAATTTACATGATGGAGAAAATTTGATCCA 1239  
 Db 1 GGTCAATCTTTGGCATTATACAGGATAAAATTTACATGATGGAGAAAATTTGATCCA 60  
 QY 1240 ACTGGGAATGTGACCAATGATGTCAGAGTTTTCACATTCATATGATGTCGGTGTG 1299  
 Db 61 ACTGGGAATGTGACCAATGATGTCAGAGTTTTCACATTCATATGATGTCGGTGTG 120  
 QY 1300 TTGACCCCTTAAGGCAAGGAGCAGTATGTCAGTGGTGGGCACCTCTGCACACATTTGTACA 1359  
 Db 121 TTGACCCCTTAAGGCAAGGAGCAGTATGTCAGTGGTGGGCACCTCTGCACACATTTGTACA 180  
 QY 1360 CTGAGGAATGCCAGTGGTCATGCTGTCATCTTTGGTCACTGCCCTCTCTATGGATAT 1419  
 Db 181 CTGAGGAATGCCAGTGGTCATGCTGTCATCTTTGGTCACTGCCCTCTCTATGGATAT 240  
 QY 1420 ATAGCAATGTGAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACCCAG 1479  
 Db 241 ATAGCAATGTGAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACCCAG 300  
 QY 1480 GGTGCCCTTTGCHAGGGGTTACGGCCATAGCAGTGTTTACACCATAGGACGGCC 1539  
 Db 301 GGTGCCCTTTGCHAGGGGTTACGGCCATAGCAGTGTTTACACCATAGGACCA-GGCC 359  
 QY 1540 CTATACCTTCATGGTGGCTACAGGCTTTTCAGTGGCAATAGTACCGGCTTTCAGATGAT 1599  
 Db 360 CTATACCTTCATGGTGGCTACAGGCTTTTCAGTGGCAATAGTACCGGCTTTCAGATGAT 419  
 QY 1600 CTCTACCGATATGATGTGGATACCCAGATGTGGACCATCTTTAAGACAGCCGATTTTC 1659  
 Db 420 CTCTACCGATATGATGTGGATACCCAGATGTGGACCATCTTTAAGACAGCCGATTTTC 479  
 QY 1660 CGTTACTTGCACACAGCTGTGATGATGAGTGGGAACCATGCTGGTGT 1706  
 Db 480 CGTTACTTGCACACAGCTGTGATGATGAGTGGGAACCATGCTGGTGT 526

RESULT 33

**BF853965**  
**LOCUS** 560 bp mRNA linear EST 16-JAN-2001  
**DEFINITION** MR2-EN0093-261200-004-h12 EN0093 Homo sapiens CDNA, mRNA sequence.  
**ACCESSION** BF853965  
**VERSION** BF853965.1 GI:12241709  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
**REFERENCE** 1 (bases 1 to 560)  
**AUTHORS** Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eumammalia; Primates; Catarrhini; Hominidae; Homo.  
 Dias Neto E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., de Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalhal, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
**TITLE** Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
**MEDLINE** 20202683  
**COMMENT** 10737800  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR2&t2=MR2-EN0093-261200-004-h12&t3=2000-12-26&t4=1)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 559.  
**FEATURES** Location/Qualifiers  
 1..560  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="EN0093"  
 /note="Organ: lung normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196, 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
**ORIGIN**  
 Query Match 8.2%; Score 353; DB 10; Length 560;  
 Best Local Similarity 99.8%; Pred. No. 4.1e-157;  
 Matches 403; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 3337 GGGCAGCGTCTCTGTGCAACACCAACGCGGCAAGTCTTTCACACCAAGGGGTC 3396  
 Db 62 GGGCAGCGTCTCTGTGCAACACCAACGCGGCAAGTCTTTCACACCAAGGGGTC 121  
 QY 3397 AAGGGGACGAGTCCAGCTATGTAGTAAATCGATACCAAGGAACCTCTCAGA 3456  
 Db 122 AAGGGGACGAGTCCAGCTATGTAGTAAATCGATACCAAGGAACCTCTCAGA 181  
 QY 3457 GGAACATGTTATTATCTCTTTCTTTTACTTACCTTACCTTTAGTCTATCCAGAA 3516  
 Db 182 GGAACATGTTATTATCTCTTTCTTTTACTTACCTTACCTTTAGTCTATCCAGAA 241  
 QY 3517 GATGATCGCTATTACAGCTATCAATTTTGTGCTACTCTCGACGAACAAACAGGAT 3576  
 Db 242 GATGATCGCTATTACAGCTATCAATTTTGTGCTACTCTCGACGAACAAACAGGAT 301  
 QY 3577 TTGGACATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGGTGCCAGT 3636

Db 302 TTGACATGTTCAATGCTCCAGATTTCAACTCAATCACTGGGCTGCCAT 361  
QY 3637 TTCTAGCTGAAACCCAGGCTGGAGAGAGATCCCTGTTGTTTCAAAAACCAACATTAAG 3696  
Db 362 TTCTAGCTGAAACCCAGGCTGGAGAGAGATCCCTGTTGTTTCAAAAACCAACATTAAG 421  
QY 3697 GAGTACAAAGATAGTTCTCTAATGAGAGTTGATTTTCGCAA 3740  
Db 422 GAGTACAAAGATAGTTCTCTAATGAGAGTTGATTTTCGCAA 465

RESULT 34  
LOCUS BG673763 390 bp mRNA linear EST 04-FEB-2002  
DEFINITION 9010 ciliated epithelial CDNA cell library Homo sapiens cDNA 5', mRNA sequence.  
ACCESSION BG673763  
VERSION BG673763.2 GI:18490081  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 390)  
AUTHORS Maiti,A.K., Jorissen,M. and Bouvagnet,P.  
TITLE Isolation, in silico characterization and chromosomal localization of a group of cDNAs from ciliated epithelial cells after in vitro ciliogenesis  
JOURNAL Genome Biol. 2 (7), RESEARCH0026 (2001)  
MEDLINE 21407920  
PUBMED 11516339  
COMMENT On May 1, 2001 this sequence, version replaced gi:13919402.  
Contact: Maiti AK  
Laboratory de Genetic Moleculaire Humaine, Faculty de Pharmacy  
University Cl. Bernard, F69373 Lyon cedex, France  
8 Avenue Rockefeller, F69373 Lyon cedex, France  
Tel: (33) 478 77 44 25  
Fax: (33) 478 77 75 68  
Email: amit.maiti@medecine.unige.ch  
Seq primer: M13 Forward and reverse.  
Location/Qualifiers  
1..390  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="Nasal biopsy"  
/cell\_type="cultured ciliated epithelial cells after in vitro Ciliogenesis"  
/dev\_stage="Adult"  
/clone\_lib="ciliated epithelial CDNA cell library"  
/note="vector: PCR2.1; cloned RT-PCR fragment from isolated total RNA"

FEATURES  
source  
1..390  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="Nasal biopsy"  
/cell\_type="cultured ciliated epithelial cells after in vitro Ciliogenesis"  
/dev\_stage="Adult"  
/clone\_lib="ciliated epithelial CDNA cell library"  
/note="vector: PCR2.1; cloned RT-PCR fragment from isolated total RNA"

ORIGIN  
Query Match 8.2%; Score 352; DB 12; Length 390;  
Best Local Similarity 100.0%; Pred. NO. 1.le-156;  
Matches 352; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 852 TCCTCACTGTACAGACAACCTGTTTTCCTCATCGAGGATCTGCAATCAAGTGATGT 911  
Db 1 TCCTCACTGTACAGACAACCTGTTTTCCTCATCGAGGATCTGCAATCAAGTGATGT 60  
QY 912 CAGAGGATGCTTCCTGCTTCTCAGACTGGCAGGGTCTCGGATGTTCAGTTCCTGTACCAGC 971  
Db 61 CAGAGGATGCTTCCTGCTTCTCAGACTGGCAGGGTCTCGGATGTTCAGTTCCTGTACCAGC 120  
QY 972 TAACAGGTCATTTTGGACTCGAGAGGAATATTTCACTTAAAGCTCCCGAGAGCATCTCA 1031  
Db 121 TAACAGGTCATTTTGGACTCGAGAGGAATATTTCACTTAAAGCTCCCGAGAGCATCTCA 180  
QY 1032 TAAAGCTGTGCTCAATGGAACATTAATGTGGGTTGTTGGAGGATATATGTTCAACCACTC 1091

Db 181 TAAAGCTGTGCTCAATGGAACATTAATGTGGGTTGTTGGAGGATATATGTTCAACCACTC 240  
QY 1092 AGATTATTAACATGTTCTTAGCGTATGACCTTCTTAGGAGTGCGTTCCACTAAACCG 1151  
Db 241 AGATTATTAACATGTTCTTAGCGTATGACCTTCTTAGGAGTGCGTTCCACTAAACCG 300  
QY 1152 TTCTGTGAACAATGTTGTTGTAGATATGTCATTTTGGCAATATATCAAG 1203  
Db 301 TTCTGTGAACAATGTTGTTGTAGATATGTCATTTTGGCAATATATCAAG 352

RESULT 35  
LOCUS BM981492/2  
DEFINITION UI-CF-EN1-adh-j-10-0-UI.s1 UI-CF-EN1 Homo sapiens cDNA clone  
ACCESSION BM981492  
VERSION BM981492.1 GI:19604039  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 681)  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
TITLE Normalization and subtraction: two approaches to facilitate gene discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
PUBMED 8895548  
COMMENT Contact: McCray, PB  
McCray Lab  
University of Iowa  
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
Tel: 319 356 4866  
Fax: 319 356 7171  
Email: paul-mccray@uiowa.edu  
Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).  
The following repetitive elements were found in this cDNA sequence: 16-237, >LIPAZ#LINE/LI (matched complement)  
Seq primer: M13 FORWARD  
POLYA=Yes.  
Location/Qualifiers  
1..681  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="UI-CF-EN1-adh-j-10-0-UI"  
/tissue\_type="Primary Lung Cystic Fibrosis Epithelial Cells"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-CF-EN1"  
/note="Organ: Lung; Vector: pTT3-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; UI-CF-EN1 is a normalized cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lemmon and Soares, Genome Research 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pTT3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CTGCTCAGGT."

TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS  
6hr to LPS 24h  
TAG LIB=UI-CF-ENL  
TAG\_SEQ=CTGCTCAGGT

## ORIGIN

Query Match 8.1%; Score 347; DB 12; Length 681;  
Best Local Similarity 99.7%; Pred. No. 3.2e-154;  
Matches 397; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3404 ACAGTCCACCTATGTGAGTAGAATAATCGATACCAAGAAACCTCTCAGAGGACAT 3463  
Db |||||  
QY 633 ACAGTCCACCTATGTGAGTAGAATAATCGATACCAAGAAACCTCTCAGAGGACAT 574  
Db |||||  
QY 3464 GTTATTATACCTCTCTTATTGACTATCAGTTTACCTTTAGTCTATCCAGGAAGATGATC 3523  
Db |||||  
QY 573 GTTATTATACCTCTCTTATTGACTATCAGTTTACCTTTAGTCTATCCAGGAAGATGATC 514  
Db |||||  
QY 3524 GCTATTACAGCTATCAATTTTGTGCTACTCTCTGACGAAACAAACAGGGATTTGGACA 3583  
Db |||||  
QY 513 GCTATTACAGCTATCAATTTTGTGCTACTCTCTGACGAAACAAACAGGGATTTGGACA 454  
Db |||||  
QY 3584 TGTTTCATCAATGCTCCAAAGAAATTTCAACCTCAACATCAGCTGGCTGCCAGTTTCTCAG 3643  
Db |||||  
QY 453 TGTTTCATCAATGCTCCAAAGAAATTTCAACCTCAACATCAGCTGGCTGCCAGTTTCTCAG 394  
Db |||||  
QY 3644 CTGAAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACA 3703  
Db |||||  
QY 393 CTGAAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACA 334  
Db |||||  
QY 3704 AAGATGTTTCTCTAAAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACA 3763  
Db |||||  
QY 333 AAGATGTTTCTCTAAAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACA 274  
Db |||||  
QY 3764 TTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 3801  
Db |||||  
QY 273 TTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 236  
Db |||||

RESULT 36  
BE259981  
LOCUS 601148539f1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:3163953 5',  
DEFINITION mRNA sequence.  
ACCESSION BE259981  
VERSION BE259981.1 GI:9130910  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 493)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: L1CM125 row: m column: 10  
High quality sequence stop: 490.  
Location/Qualifiers  
1. .493  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3163953"  
/tissue\_type="neuroblastoma"  
/lab\_host="DH10B (phage-resistant)"

/clone lib="NIH\_MGC\_19"  
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:  
ECORI; cDNA made by oligo-dT priming. Directionally  
cloned into ECORI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 7.8%; Score 335; DB 10; Length 493;  
Best Local Similarity 100.0%; Pred. No. 1.7e-148;  
Matches 335; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1354 GTTACACTGAAGAATGGCCGAGTGTCTCATCTTTGGTCACTGCCCTCTCTAT 1413  
Db |||||  
QY 158 GTTACACTGAAGAATGGCCGAGTGTCTCATCTTTGGTCACTGCCCTCTCTAT 217  
Db |||||  
QY 1414 GGAATATAAGCAATGTGCAGGAATATGATTTGGATAAGAAACACATGAGATATTACAC 1473  
Db |||||  
QY 218 GGAATATAAGCAATGTGCAGGAATATGATTTGGATAAGAAACACATGAGATATTACAC 277  
Db |||||  
QY 1474 ACCCAGGCTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTCAGCAGCATAGGACC 1533  
Db |||||  
QY 278 ACCCAGGCTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTCAGCAGCATAGGACC 337  
Db |||||  
QY 1534 AGGGCCCTTATACGTTTCATGTTGGTGTACAAAGGCTTTTCAGTGCCCAATAAGTACCGGCTTGCA 1593  
Db |||||  
QY 338 AGGGCCCTTATACGTTTCATGTTGGTGTACAAAGGCTTTTCAGTGCCCAATAAGTACCGGCTTGCA 397  
Db |||||  
QY 1594 GATGATCTCTACCGATATGATGTGATACCCAGATGGACCATCTTAAAGGACAGCGGA 1653  
Db |||||  
QY 398 GATGATCTCTACCGATATGATGTGATACCCAGATGGACCATCTTAAAGGACAGCGGA 457  
Db |||||  
QY 1654 TTTTTCCTTACTTGTGCACACAGCTGTGATAGTGAG 1688  
Db |||||  
QY 458 TTTTTCCTTACTTGTGCACACAGCTGTGATAGTGAG 492  
Db |||||

## RESULT 37

BE259981  
LOCUS 602853292f1 NIH\_MGC\_10 Homo sapiens cDNA clone IMAGE:4994735 5',  
DEFINITION mRNA sequence.  
ACCESSION BE259981  
VERSION BE259981.1 GI:14507662  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 706)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: Incyte Genomics, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: L1AM1017 row: g column: 24  
High quality sequence stop: 676.  
Location/Qualifiers  
1. .706  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4994735"

## FEATURES

## source



Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matekuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV1&t2=QV1-ET0181-021200-519-d10&t3=2000-12-02&t4=1>)

Seq primer: puc 18 forward

High quality sequence start: 11

High quality sequence stop: 399.

Location/Qualifiers

1. .399

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="ET0181"

/note="Organ: lung tumor; Vector: puc18; Site: 1: Smal;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORSTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the pUC 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

# FEATURES

source

## FEATURES

source

1. .671

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:2424606"

/tissue\_type="well-differentiated endometrial

adenocarcinoma, 7 pooled tumors"

/lab\_host="DH10B"

/clone\_lib="NCI CGAP Util"

/note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.

Average insert size 1.75 kb. Life Technologies catalog #:

11538-014"

## ORIGIN

Query Match

Best Local Similarity 7.0%; Score 300; DB 9; Length 671;

Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3502 AGTCATATCCAGGAGATGATCGCTATTACACAGTATCAATTTTGGGCTACTCTGAC 3561

DB 517 AGTCATATCCAGGAGATGATCGCTATTACACAGTATCAATTTTGGGCTACTCTGAC 458

QY 3562 GACAAACAGGATTTGGACATGTTTCATCAATGCTCCAGAAATTTCACTCAACATC 3621

DB 457 GAACAAACAGGATTTGGACATGTTTCATCAATGCTCCAGAAATTTCACTCAACATC 398

QY 3622 ACCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTGTGTTTCA 3681

DB 397 ACCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTGTGTTTCA 338

QY 3682 AAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGAGAGATTTGATTTCCGAAC 3741

DB 337 AAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGAGAGATTTGATTTCCGAAC 278

QY 3742 CACCAATATACATCTTCTTTGTTTATGTCAGTAAATTTCACTGCCCATCAAAATTCAG 3801

DB 277 CACCAATATACATCTTCTTTGTTTATGTCAGTAAATTTCACTGCCCATCAAAATTCAG 218

RESULT 41

AI372809/c

LOCUS

DEFINITION

AI372809

ESL175353 Infant brain, Bento Soares Homo sapiens cDNA clone

linear EST 12-JAN-1999

RESULT 40

AI1818764/c

LOCUS

AI1818764

671 bp mRNA

linear EST 07-MAR-2000



```

FUHIBU1, mRNA sequence.
ACCESSION AI372809
VERSION AI372809.1 GI:4152675
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 383)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geohagen,N.S.,
Glodek,A., Gnehm,C.B., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
TITLE Initial assessment of human gene diversity and expression patterns
based upon 93 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
PUBMED 7566098
COMMENT Other ESTs: EST175328 EST175329 EST175330 EST175331 EST175332
EST175333 EST175334 EST175335 EST175336 EST175337 EST175338
EST175339 EST17534
Contact: HGI (Human Gene Index)
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-0200
Fax: (301)-838-0208
Email: hgi@tigr.org.
Location/Qualifiers
FEATURES
source
1..383
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FUHIBU1"
/lab_host="E. coli DH5-alpha"
/clone_lib="infant brain, Bento Soares"
/notes="Vector: EA, M13-derived; Site 1: HindIII; Site 2:
NotI; The infant brain library, constructed by Bento
Soares, Columbia University, was oligo-(dT) primed and
directionally cloned into an M13-derived plasmid using
total brain mRNA from a 72-day old human female afflicted
with spinal muscular atrophy."
ORIGIN
Query Match 7.0%; Score 299; DB 9; Length 383;
Best Local Similarity 100.0%; Pred. No. 2.7e-131;
Matches 299; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2462 CCCTTTTGGCTTCCTTACACCCAGAGAGGTAGATTTGCTTAAAGCAGCTCGAA 2521
DB 299 CCCTTTTGGCTTCCTTACACCCAGAGAGGTAGATTTGCTTAAAGCAGCTCGAA 240
QY 2522 TAATGCAGTCATCTCAGAGCATGTCCCAAGCTCACCTTTAACCCCATGGGTGGCTTCGGA 2581
DB 239 TAATGCAGTCATCTCAGAGCATGTCCCAAGCTCACCTTTAACCCCATGGGTGGCTTCGGA 180
QY 2582 AGATCAATGTGTCTACTGTGCTGGGAAGATATGTCCCATTTACAAATAGTTTACTAC 2641
DB 179 AGATCAATGTGTCTACTGTGCTGGGAAGATATGTCCCATTTACAAATAGTTTACTAC 120
QY 2642 AGTGGATGCCGTCTGAGCCAGTCATGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTA 2701

```

```

119 AGTGGATCCGCTCTGAGCCCGAGTGTGGATTCTGGAAATTTATCAGAACCCAGTA 60
QY 2702 CTCGGGGACTGAAGGCTGCACCTCATCAACCCCACTCAATGTAAGTGTCTGTGAAAGG 2760
DB 59 CTCGGGGACTGAAGGCTGCACCTCATCAACCCCACTCAATGTAAGTGTCTGTGAAAGG 1
RESULT 42
AW579586
LOCUS MRO-HT0209-100100-103-e09 HT0209 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW579586
VERSION AW579586.1 GI:7254635
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 427)
HCGP http://www.ludwig.org.br/ORESTES.
The FAPESP/LICR Human Cancer Genome Project
Unpublished (1999)
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MRO&t2=NR0-HT0209-
100100-103-e09&t3=2000-01-10&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 427.
Location/Qualifiers
FEATURES
source
1..427
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="HT0209"
/clone_stage="Adult"
/notes="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
ORIGIN
Query Match 6.9%; Score 298; DB 10; Length 427;
Best Local Similarity 99.7%; Pred. No. 8.4e-131;
Matches 348; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1299 GTTGACCCCTTAAGGCAAGGAGCAGTATGCAGTGTGTGGCAGCTCGACATTTGTAC 1358
DB 63 GTTGACCCCTTAAGGCAAGGAGCAGTATGCAGTGTGTGGCAGCTCGACATTTGTAC 122
QY 1359 ACTGAGATGCGCCAGTGGTTCATCTTTGTCATCTTTGTCATCTTCCTCTATGGATA 1418
DB 123 ACTGAGATGCGCCAGTGGTTCATCTTTGTCATCTTTGTCATCTTCCTCTATGGATA 182
QY 1419 TATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGAGTATATTACACCCA 1478
DB 183 TATAAGCAATGTGCAGGAATATGATTTGGATAAGAACACATGAGTATATTACACCCA 242
QY 1479 GGGTGCCTTGTGGAAGGGGGTTACGGCCATAGCAGTGTTCAGCACCATAGACAGGCGC 1538
DB 243 GGGTGCCTTGTGGAAGGGGGTTACGGCCATAGCAGTGTTCAGCACCATAGACAGGCGG 302

```

```

QY 1539 CCTATACGTTTCATGTCGTCACAAAGGCTTTTCAGTGCACCAATAGTACCGGCTTGCAGATGA 1598
DB 303 CCTATACGTTTCATGTCGTCACAAAGGCTTTTCAGTGCACCAATAGTACCGGCTTGCAGATGA 362
QY 1599 TCTCTACCGATATGATGTCGATACCCAGATGTCGACCATCTTAAAGAC 1647
DB 363 TCTCTACCGATATGATGTCGATACCCAGATGTCGACCATCTTAAAGAC 411

RESULT 43
LOCUS AI372810/c 320 bp mRNA linear EST 12-JAN-1999
DEFINITION EST175354 Infant brain, Bento Soares Homo sapiens cDNA clone
FUHIBU1, mRNA sequence.
ACCESSION AI372810
VERSION AI372810.1 GI:4152676
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 320)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geonhagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaro,S.M., Merrick,J.M.,
Moreno-Palancques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,S.J.,
Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Meissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M., and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
96026280
7566098
COMMENT Other ESTs: EST175328 EST175329 EST175330 EST175331 EST175332
EST175333 EST175334 EST175335 EST175336 EST175337 EST175338
EST175339 EST17534
Contact: HGI (Human Gene Index)
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-0200
Fax: (301)-838-0208
Email: hgi@tigr.org.

FEATURES
Location/Qualifiers
1..320
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FUHIBU1"
/lab_host="E. coli DH5-alpha"
/clone_lib="Infant brain, Bento Soares"
Note=Vector: BA, M13-derived; Site_1: HindIII; Site_2:
NotI; The infant brain library, constructed by Bento
Soares, Columbia University, was oligo-(dT) primed and
directionally cloned into an M13-derived plasmid using
total brain mRNA from a 72-day old human female afflicted
with spinal muscular atrophy."

ORIGIN
Query Match 6.6%; Score 283; DB 9; Length 320;
Best Local Similarity 100.0%; Pred. No. 1.2e-123;
Matches 283; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2761 CTTGCAACCAACAGTCGTTAAGCAGTCCGGACACCATGTCCTTGAGACAGCATGTGGA 2820
DB 334 CTTGCAACCAACAGTCGTTAAGCAGTCCGGACACCATGTCCTTGAGACAGCATGTGGA 275
QY 2821 GATTGCAACCAACAGTCGTTAAGCAGTCCGGACACCATGTCCTTGAGACAGCATGTGGA 2880

```

```

QY 2462 CCCTTTTGCTTCTCTTCAAAACCCAGAAAGGTAGAAATTTGTCTTAAAGCAGCTGCAG 2521
DB 283 CCCTTTTGCTTCTCTTCAAAACCCAGAAAGGTAGAAATTTGTCTTAAAGCAGCTGCAG 224
QY 2522 TAATGCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCGTGGTGGCTTTCGGA 2581
DB 223 TAATGCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCGTGGTGGCTTTCGGA 164
QY 2582 AGATCAATGTGTCTTACTGTGCTGGGAAGATATGTCCCAATTTACAATATGTTTACTAC 2641
DB 163 AGATCAATGTGTCTTACTGTGCTGGGAAGATATGTCCCAATTTACAATATGTTTACTAC 104
QY 2642 AGTGAATGCTGTGAGCCAGTGTGTCGATTTCTGTGGAATTTATCAGAAACCCAGTA 2701
DB 103 AGTGAATGCTGTGAGCCAGTGTGTCGATTTCTGTGGAATTTATCAGAAACCCAGTA 44
QY 2702 CTCGGGGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGG 2744
DB 43 CTCGGGGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGG 1

RESULT 44
LOCUS AA350293/c 400 bp mRNA linear EST 21-APR-1997
DEFINITION EST57512 Infant brain Homo sapiens cDNA 5' end, mRNA sequence.
ACCESSION AA350293
VERSION AA350293.1 GI:2002620
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 400)
Adams,M.D., Kerlavage,A.R., Fields,C. and Venter,J.C.
3,400 expressed sequence tags identify diversity of transcripts
from human brain
Nat. Genet. 4, 256-267 (1993)
93364420
8358434
COMMENT Other ESTs: EST57511, THC168226
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.

FEATURES
Location/Qualifiers
1..400
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):151062"
/db_xref="taxon:9606"
/sex="female"
/dev stage="infant"
/clone_lib="Infant brain"
Note=Organ: brain; Vector: lafmid BA; Site_1: HindIII;
Site_2: NotI"

ORIGIN
Query Match 6.6%; Score 283; DB 9; Length 400;
Best Local Similarity 99.7%; Pred. No. 1.3e-123;
Matches 333; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2761 CTTGCAACCAACAGTCGTTAAGCAGTCCGGACACCATGTCCTTGAGACAGCATGTGGA 2820
DB 334 CTTGCAACCAACAGTCGTTAAGCAGTCCGGACACCATGTCCTTGAGACAGCATGTGGA 275
QY 2821 GATTGCAACCAACAGTCGTTAAGCAGTCCGGACACCATGTCCTTGAGACAGCATGTGGA 2880

```

```
Db 274 GATTGCACGCGNAGCTCTGAGTGTCATGCTGGTGAGCAACATGAAGCAGTGTGTGGAC 215
Qy 2881 TCCAATGCCATATGGGCTCTCCCTTTGGCCAGTGTATGGAATGGTATACATGAGC 2940
Db 214 TCCAATGCCATATGGGCTCTCCCTTTGGCCAGTGTATGGAATGGTATACATGAGC 155
Qy 2941 ACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCA 3000
Db 154 ACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCA 95
Qy 3001 GGCTGTGGCTGGTGTACTGATCCAGCAATCTGGCAAGGAAATGTCATAGAGGTTCC 3060
Db 94 GGCTGTGGCTGGTGTACTGATCCAGCAATCTGGCAAGGAAATGTCATAGAGGTTCC 35
Qy 3061 TATAAGGACGAGTGAAGTGCCTTGCAGGCC 3094
Db 34 TATAAGGACGAGTGAAGTGCCTTGCAGGCC 1

RESULT 45
BF853021/c
LOCUS BF853021 483 bp mRNA linear EST 16-JAN-2001
DEFINITION MR2-EN0093-191200-001-905 EN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF853021
VERSION BF853021.1 GI:12240869
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 483)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
PUBLISHED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-
191200-001-G05&t3=2000-12-19&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 483.
Location/Qualifiers
1..483
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0093"
/note="Organ: lung normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

FEATURES
source
```

ORIGIN

```
Query Match 6.6%; Score 281; DB 10; Length 483;
Best Local Similarity 99.5%; Pred. No. 1.2e-122;
Matches 381; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 3419 GTGAGGTAGAAAATCGATACCAAGAAACCCCTCTCAGAGGAACATGTTATTATATCTTC 3478
Db 448 GTGAGGTAGAAAATCGATACCAAGAAACCCCTCTCAGAGGAACATGTTATTATATCTTC 389
Qy 3479 TTATTGACTATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTATACAGCTA 3538
Db 388 TTATTGACTATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTATACAGCTA 329
Qy 3539 TCAATTTTGTGGTACTCTCTGACGAACAAACAGGATTTGGACATGTTCAATGCTT 3598
Db 328 TCAATTTTGTGGTACTCTCTGACGAACAAACAGGATTTGGACATGTTCAATGCTT 269
Qy 3599 CCAAGAATTTCAACCTCAACATCACCTGGGTGCCAGTTTCTCAGCTGGAACCCAGGCTG 3658
Db 268 CCAAGAATTTCAACCTCAACATCACCTGGGTGCCAGTTTCTCAGCTGGAACCCAGGCTG 209
Qy 3659 GAGAAGAGATGCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATGTTCTCTA 3718
Db 208 GAGAAGAGATGCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATGTTCTCTA 149
Qy 3719 ATGAGAAGTTTCATTTTGGCAACCAACCAATATACACTTTCTTTTATGTCAGTAAT 3778
Db 148 ATGAGAAGTTTCATTTTGGCAACCAACCAATATACACTTTCTTTTATGTCAGTAAT 89
Qy 3779 TCACCTGGCCCATCAAAATTCAG 3801
Db 88 TCACCTGGCCCATCAAAATTCAG 66

RESULT 46
BG015062/c
LOCUS BG015062 281 bp mRNA linear EST 24-JAN-2001
DEFINITION QV2-GN0208-221200-579-a02 GN0208 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG015062
VERSION BG015062.1 GI:12466889
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 281)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
PUBLISHED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV2&t2=QV2-GN0208-
221200-579-a02&t3=2000-12-22&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 281.
Location/Qualifiers
1..281
/organism="Homo sapiens"

FEATURES
source
```

```
/mol_type="mrna"  
/db_xref="taxon:9606"  
/dev_stage="Adult"  
/clone_lib="GNO208"  
/note="Organ: placenta normal; Vector: puc18; Site 1:  
SmaI; Site 2: SmaI; A Mini-library was made by cloning  
products derived from ORESTES PCR (U.S. Letters Patent  
application No. 196,716 - Ludwig Institute for Cancer  
Research) profiles into the pUC18 vector. Reverse  
transcription of tissue mRNA and cDNA amplification were  
performed under low stringency conditions."
```

## ORIGIN

```
Query Match 6.3%; Score 270; DB 10; Length 281;  
Best Local Similarity 100.0%; Pred. No. 2e-117;  
Matches 270; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1760 CTTGAGATTTGATGGCCCTATGACATTTGCTGTGACCGCTGTGAGTCTTCCAGACCTG 1819  
Db CTTGAGATTTGATGGCCCTATGACATTTGCTGTGACCGCTGTGAGTCTTCCAGACCTG 222  
  
QY 1820 ATCTCCACCATGATGTCAACAGATTGGCCATTGACGAGTCTTACACAAAGCAGCACCATTG 1879  
Db ATCTCCACCATGATGTCAACAGATTGGCCATTGACGAGTCTTACACAAAGCAGCACCATTG 162  
  
QY 1880 ATGTGTCGGTGGTTCAATAGTCTCTCTCCAGCAGATCTTGGTATTCACTCGGAAC 1939  
Db ATGTGTCGGTGGTTCAATAGTCTCTCTCCAGCAGATCTTGGTATTCACTCGGAAC 102  
  
QY 1940 AGTGTGATGCCATCGGAGTGAACCGCTGTTTGTAGCAGGAGCCTGTTATTCGGTGTG 1999  
Db AGTGTGATGCCATCGGAGTGAACCGCTGTTTGTAGCAGGAGCCTGTTATTCGGTGTG 42  
  
QY 2000 TGTGGAACACAGGGTGGCTCTCAGTGTATCT 2029  
Db TGTGGAACACAGGGTGGCTCTCAGTGTATCT 12
```

```
RESULT 47  
AA350292  
LOCUS EST57511 Infant brain Homo sapiens cDNA 3' end, mRNA sequence.  
ACCESSION AA350292  
VERSION AA350292.1 GI:2002619  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Adams, M.D., Kerlavage, A.R., Fields, C. and Venter, J.C.  
3,400 expressed sequence tags identify diversity of transcripts  
from human brain  
Nat. Genet. 4, 256-267 (1993)  
93364420  
8358434  
Other ESTs: EST57512 THC168885  
Contact: Kerlavage, AR  
Bioinformatics  
The Institute for Genomic Research  
9712 Medical Center Drive, Rockville, MD 20850 USA  
Tel: 3018699056  
Fax: 3018699423  
Email: arkerlav@tigr.org  
For clone availability, additional sequence and expression  
information related to this EST, please check the TIGR Human Gene  
Index (http://www.tigr.org/tdb/hgi/hgi.html)  
Seq primer: M13-21.  
Location/Qualifiers  
1. .396  
/organism="Homo sapiens"  
/mol_type="mrna"  
/db_xref="ATCC (inhost):151062"
```

```
/db_xref="taxon:9606"  
/sex="female"  
/dev_stage="infant"  
/clone_lib="Infant brain"  
/note="Organ: brain; Vector: lafmid BA; Site 1: HindIII;  
Site 2: NotI"
```

## ORIGIN

```
Query Match 6.3%; Score 270; DB 9; Length 396;  
Best Local Similarity 99.5%; Pred. No. 2.1e-117;  
Matches 370; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1438 TATGATTTGGATAAGAACACATGGAGTATATTACACCCAGGGTGCCCTTGTGCAAGG 1497  
Db TATGATTTGGATAAGAACACATGGAGTATATTACACCCAGGGTGCCCTTGTGCAAGG 60  
  
QY 1498 GGTTCAGGCCCATAGCAGTGTTTACACCATAGGACCGCCCTATACGTTTCATGGTGGC 1557  
Db GGTTCAGGCCCATAGCAGTGTTTACACCATAGGACCGCCCTATACGTTTCATGGTGGC 120  
  
QY 1558 TACAAGGGCTTTTCAGTGCCTCAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATG 1617  
Db TACAAGGGCTTTTCAGTGCCTCAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATG 180  
  
QY 1618 GATACCCAGATGTGACCAATTCTTAAAGACAGCCGATTTTCCGTTACTTGCACACAGCT 1677  
Db GATACCCAGATGTGACCAATTCTTAAAGACAGCCGATTTTCCGTTACTTGCACACAGCT 240  
  
QY 1678 GTGATAGTGAAGTGAACCATGCTGCTGTTGGGGGAAACACACAAATGACACATCTATG 1737  
Db GTGATAGTGAAGTGAACCATGCTGCTGTTGGGGGAAACACACAAATGACACATCTATG 300  
  
QY 1738 AGCCATGGCGCAAAATGCTTCTTCAGATTTCAATGCGCTATGACATTCCTGTGACCGC 1797  
Db AGCCATGGCGCAAAATGCTTCTTCAGATTTCAATGCGCTATGACATTCCTGTGACCGC 360  
  
QY 1798 TGGTCAGTGCCT 1809  
Db TGGTCAGTGCCT 372
```

```
RESULT 48  
BE709733/c  
LOCUS IL3-HT0618-150600-177-All HT0618 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BE709733  
ACCESSION BE709733  
VERSION BE709733.1 GI:10097998  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
MEDLINE  
10737800  
PUBMED  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome
```

Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=IL3-HT0618-150  
600-177-A11&t3=2000-05-15&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 22  
High quality sequence stop: 447.  
Location/Qualifiers

## FEATURES

source  
1. 524  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="HT0618"

/note="Organ: head neck; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

## ORIGIN

Query Match 6.3%; Score 269; DB 10; Length 524;  
Best Local Similarity 99.7%; Pred. No. 6.8e-117; Indels 0; Gaps 0;  
Matches 319; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 481 CTCATTGAAGGACAGCCAAATAGAAATAGAGATTCGTTTCAATCATTTTGTACAGAG 540  
Db 524 CTCATTGAAGGACAGCCAAATAGAAATAGAGATTCGTTTCAATCATTTTGTACAGAG 465  
QY 541 TGTAGTTGGGACCAATTTATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCT 600  
Db 464 TGTAGTTGGGACCAATTTATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCT 405  
QY 601 GCATTTAGTGGCTTCATTTGTTCTTGAGAGATGGCAATGAGACTGTCCTGAGTTGTT 660  
Db 404 GCATTTAGTGGCTTCATTTGTTCTTGAGAGATGGCAATGAGACTGTCCTGAGTTGTT 345  
QY 661 GCCACATCAGTTATGCTGTCGCAATTTTATGATGCTGCTATATATTTGACTGGA 720  
Db 344 GCCACATCAGTTATGCTGTCGCAATTTTATGATGCTGCTATATATTTGACTGGA 285  
QY 721 TTTAATATTACTACAGTTTGTATGTTGTCCTCAATTAATGCTCAGGCCGAGGAGAGTGT 780  
Db 284 TTTAATATTACTACAGTTTGTATGTTGTCCTCAATTAATGCTCAGGCCGAGGAGAGTGT 225  
QY 781 AAGATCAGTAATACGCGGA 800  
Db 224 AAGATCAGTAATACGCGGA 205

## RESULT 49

BU531905 778 bp mRNA linear EST 13-SEP-2002  
LOCUS  
DEFINITION AGENCOURT 10205588 NIH MGC 126 Homo sapiens cDNA clone  
IMAGE:6556819 5', mRNA sequence.

ACCESSION BU531905.1 GI:22842346

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 778)

NIH-MGC http://mgi.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: NCI

CNA Library Preparation: Michael Brownstein Laboratory

CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov

Plate: LLCW2723 row: n column: 19

High quality sequence stop: 480.

## FEATURES

source  
1. 778  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6556819"  
/tissue\_type="mixed (pool of 40 RNAs)"  
/lab\_host="DH10B (T1-phage-resistant)"

/clone\_lib="NIH\_MGC\_126"  
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);  
Site 2: SfiI (ggcgctcgcc); Double-stranded cDNA was  
prepared from a pool of 40 cell line polyA+ RNAs (bladder  
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -  
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,  
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell -  
5.2%, ovary - 4%, pharynx - 2.3%, prostate - 4.3%,  
salivary gland - 1.3%, and skin - 2.3%). 5' and 3'

adaptors were used in cloning as follows:

5'-ATTCTAGAGCGCGCGCGACATG-DT(30)NN-3'. Full-length  
5'-AAGCAGTGTATCAACGAGAGTGGCATTACGGCGG-3' and  
enriched library was constructed using the Clontech  
Creator SMART kit and size-selected to contain the 0.5-1  
kb size fraction (other fractions present in NIH\_MGC\_127  
and NIH\_MGC\_128). Library created in the laboratory of T.  
Usdin, M.D., Ph.D. (NIMH, NIH). Note: this is a NIH\_MGC  
Library."

## ORIGIN

Query Match 5.9%; Score 253; DB 13; Length 778;  
Best Local Similarity 100.0%; Pred. No. 3.4e-109; Indels 0; Gaps 0;  
Matches 253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3549 GGTCTACTCTGACGACAAACAGGAGTTTGACATGTTTCATCAATGCTCCCAAGATTT 3608  
Db 3 GGTCTACTCTGACGACAAACAGGAGTTTGACATGTTTCATCAATGCTCCCAAGATTT 62  
QY 3609 CAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGTGAACCCAGGCTGGAGAGAGAT 3668  
Db 63 CAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGTGAACCCAGGCTGGAGAGAGAT 122  
QY 3669 GCCTGTGTTTCAAAACCAACATTAAAGAGTACAAAGATGTTCTCTTAATGAGAAGTT 3728  
Db 123 GCCTGTGTTTCAAAACCAACATTAAAGAGTACAAAGATGTTCTCTTAATGAGAAGTT 182  
QY 3729 TGATTTTCGCAACCCACCAATATCACTTTCTTTTATGTCAGTAATTTCCCTGGCC 3788  
Db 183 TGATTTTCGCAACCCACCAATATCACTTTCTTTTATGTCAGTAATTTCCCTGGCC 242  
QY 3789 CATCAAAATTCAG 3801  
Db 243 CATCAAAATTCAG 255

## RESULT 50

AI818750/c

LOCUS

DEFINITION

AI818750.1

VERSION

KEYWORDS

SOURCE

ORGANISM

AI818750

EST.

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 589)

NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CCAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 2056 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 399.  
Location/Qualifiers  
source  
1. 589  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2424580"  
/tissue\_type="well-differentiated endometrial  
adenocarcinoma, 7 pooled tumors"  
/lab\_hosts="DH10B"  
/clone\_lib="NCI CGAP Utl1"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"

## ORIGIN

Query Match 5.8%; Score 249; DB 9; Length 589;  
Best Local Similarity 99.7%; Pred. No. 2.6e-107;  
Matches 299; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3502 AGTCTATCCGAGGAGATGCGTATACAGCTATCAATTTGGGCTACTCTGAC 3561  
DB 517 AGTCTATCCGAGGAGATGCGTATACAGCTATCAATTTGGGCTACTCTGAC 458  
QY 3562 GAAACAAACAGGGATTGGACATGTTTCATCAATCCCTCCAGAAATTTCAACCTCAACATC 3621  
DB 457 GAAACAAACAGGGATTGGACATGTTTCATCAATCCCTCCAGAAATTTCAACCTCAACATC 398  
QY 3622 ACCTGGGCTGCCATTTCTCAGCTGGAACCCAGGCTCGAAGAGATGCGCTGTGTTTCA 3681  
DB 397 ACCTGGGCTGCCATTTCTCAGCTGGAACCCAGGCTCGAAGAGATGCGCTGTGTTTCA 338  
QY 3682 AAAACCAACATTAAGGAGTACAAAGATGTTTCTCTAATCAGAGATTGATTTTCGCAAC 3741  
DB 337 AAAACCAACATTAAGGAGTACAAAGATGTTTCTCTAATCAGAGATTGATTTTCGCAAC 278  
QY 3742 CACCCAAATACATCTTTCTTTGTTTATGTCAGTAATTTCACTGGCCCATCAAAATTCAG 3801  
DB 277 CACCCAAATACATCTTTCTTTGTTTATGTCAGTAATTTCACTGGCCCATCAAAATTCAG 218

RESULT 51  
BF928756  
LOCUS IL5-NT0228-061200-308-f12 NT0228 Homo sapiens cdna, mRNA linear EST 19-JAN-2001  
DEFINITION  
ACCESSION BF928756  
VERSION BF928756.1 GI:12326884  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 313)  
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., D.H.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,

TITLE O'Hare, M.J., Soares, F., Brentani, R.R.; Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
JOURNAL Shotgun sequencing of the human transcriptome with ORF expressed  
MEDLINE sequence tags  
20202663  
10737800  
COMMENT Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l1=IL5&l2=IL5-NT0228-  
061200-308-f12&t3=2000-12-06&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 313.  
Location/Qualifiers  
source  
1. 313  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="NT0228"  
/note="Organ: nervous tumor; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

## FEATURES

Query Match 5.7%; Score 243; DB 10; Length 313;  
Best Local Similarity 99.7%; Pred. No. 1.7e-104;  
Matches 293; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3307 GGAGGGAATGTGACCCATGCAAGTGCATGGGACGCGTCTCTGCAACACCAACACG 3366  
DB 7 GGAGGGAATGTGACCCATGCAAGTGCATGGGACGCGTCTCTGCAACACCAACACG 66  
QY 3367 GGCAAGTCTCTTGTGACCAACCAAGGCGTCAAGGGGACGAGTCCAGCTATGTGAGTA 3426  
DB 67 GGCAAGTCTCTTGTGACCAACCAAGGCGTCAAGGGGACGAGTCCAGCTATGTGAGTA 126  
QY 3427 GAAATCGATACCAAGGAACCTCTCAGAGAACATGTTATTATCTCTTCTTATTGAC 3486  
DB 127 GAAATCGATACCAAGGAACCTCTCAGAGAACATGTTATTATCTCTTCTTATTGAC 186  
QY 3487 TATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTT 3546  
DB 197 TATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTT 246  
QY 3547 GTGGCTACTCTGACGAACAAACAGGATTGGACATGTTTCATCATGCTCTCC 3600  
DB 247 GTGGCTACTCTGACGAACAAACAGGATTGGACATGTTTCATCATGCTCTCC 300

## ORIGIN

Query Match 5.7%; Score 243; DB 10; Length 313;  
Best Local Similarity 99.7%; Pred. No. 1.7e-104;  
Matches 293; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 3307 GGAGGGAATGTGACCCATGCAAGTGCATGGGACGCGTCTCTGCAACACCAACACG 3366  
DB 7 GGAGGGAATGTGACCCATGCAAGTGCATGGGACGCGTCTCTGCAACACCAACACG 66  
QY 3367 GGCAAGTCTCTTGTGACCAACCAAGGCGTCAAGGGGACGAGTCCAGCTATGTGAGTA 3426  
DB 67 GGCAAGTCTCTTGTGACCAACCAAGGCGTCAAGGGGACGAGTCCAGCTATGTGAGTA 126  
QY 3427 GAAATCGATACCAAGGAACCTCTCAGAGAACATGTTATTATCTCTTCTTATTGAC 3486  
DB 127 GAAATCGATACCAAGGAACCTCTCAGAGAACATGTTATTATCTCTTCTTATTGAC 186  
QY 3487 TATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTT 3546  
DB 197 TATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTT 246  
QY 3547 GTGGCTACTCTGACGAACAAACAGGATTGGACATGTTTCATCATGCTCTCC 3600  
DB 247 GTGGCTACTCTGACGAACAAACAGGATTGGACATGTTTCATCATGCTCTCC 300  
RESULT 52  
R87660  
LOCUS yp99g01.s1 Soares fetal liver spleen INFLS Homo sapiens cdna clone  
DEFINITION IMAGE:194640 3', mRNA sequence.  
ACCESSION R87660  
VERSION R87660.1 GI:946473  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE AUTHORS	TITLE JOURNAL COMMENT	FEATURES source	ORIGIN
Mammalia; Eutharia; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 416) Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevasakis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.	The WashU-Merck EST Project Unpublished (1995) Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Insert Size: 1984 High quality sequence stops: 273 Source: IMAGE Consortium, LML This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Insert Length: 1984 Std Error: 0.00 Seq primer: Promega -21m3 High quality sequence stops: 273. Location/Qualifiers 1. 416 /organism="Homo sapiens" /mol type="mRNA" /db xref="GDB:3763690" /db xref="taxon:9606" /clone="IMAGE:194640" /sex="male" /dev stage="20 week-post conception fetus" /lab host="DH10B (ampicillin resistant)" /clone lib="Soares fetal liver spleen INFLS" /notes="Organ: Liver and Spleen; Vector: pTV73D (Pharmacia) with a modified polylinker, Site_1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' - AACTGGAGAATAATAAGACCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pTV73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."	Query Match 5.6%; Score 241; DB 14; Length 416; Best Local Similarity 100.0%; Pred. No. 1.6e-103; Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	QY 2257 ATGTACTACTGTACACAGAGACAGCTGCAGGAGCTGTCCCTGGACACAGACTGCCAG 2316 Db 1 ATGTACTACTGTACACAGAGACAGCTGCAGGAGCTGTCCCTGGACACAGACTGCCAG 60
		QY 2317 TGGGAGCCCGGAATCAGGAGTGCATTGTCCTGCCGAAATATCTGTGCATTGGCTGG 2376 Db 61 TGGGAGCCCGGAATCAGGAGTGCATTGTCCTGCCGAAATATCTGTGCATTGGCTGG 120	
		QY 2377 CATTTGGTTGGAACACTCATGTTTGAAATTTACTACTGCCAAGAGGAATTTAGCAATGCT 2436 Db 121 CATTTGGTTGGAACACTCATGTTTGAAATTTACTACTGCCAAGAGGAATTTAGCAATGCT 180	
		QY 2437 AAATTGTTCTGTAGGAACCAATGCCCTTTTGGCTTCTCTTCAACCCAGAGAAGGTA 2496 Db 181 AAATTGTTCTGTAGGAACCAATGCCCTTTTGGCTTCTCTTACACCCAGAGAAGGTA 240	
		QY 2497 G 2497 Db 241 G 241	

RESULT 53 .			
BE671400/c			
LOCUS	BE671400	445 bp	linear
		mRNA	EST 08-SEP-2000

DEFINITION	7e50d11.x1 NCI_CGAP Lu24 Homo sapiens cDNA clone IMAGE:3285909 3', similar to TR:060295 060295 KIA0548 PROTEIN ; contains 11.b1 L1 repetitive element ;, mRNA sequence.
ACCESSION	BE671400
KEYWORDS	EST.
SOURCE	BE671400.1 GI:10031941
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1 (bases 1 to 445)
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL	Unpublished (1997)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgraphs-@email.nih.gov Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: M. Bento Soares, Ph.D. cDNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: infoimage.llnl.gov Seq primer: -40UP from Gibco High quality sequence stop: 397. Location/Qualifiers
FEATURES	1..445 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:3285909" /tissue_type="carcinoid" /lab_host="DH10B" /clone_lib="NCI_CGAP Lu24" /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonids 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."
ORIGIN	5.2%; Score 225; DB 10; Length 445; Query Match 100.0%; Pred.No. 7.7e-96; Best Local Similarity 0; Mismatches 0; Indels 0; Gaps 0; Matches 225; Conservative 0
QY	3577 TTGGACATGTTTCATCAATGCCTCCAGAATTTTCAACCTCAACATCACTGGGTGCCAGT 3636
Db	445 TTGGACATGTTTCATCAATGCCTCCAGAATTTTCAACCTCAACATCACTGGGTGCCAGT 386
QY	3637 TTCTCAGCTGGAAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAA 3696
Db	385 TTCTCAGCTGGAAACCCAGGCTGGAGAAGAGATGCCTGTTGTTTCAAAAACCAACATTAA 326
QY	3697 GAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGAATTTTCGCAACCAACCCAAATATCACT 3756
Db	325 GAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGAATTTTCGCAACCAACCCAAATATCACT 266
QY	3757 TTCTTTGTTTATGTCAGTAAATTTTCACCTGGCCCATCAAAATTGAG 3801
Db	265 TTCTTTGTTTATGTCAGTAAATTTTCACCTGGCCCATCAAAATTGAG 221

RESULT 54	AI076479/c	464 bp	mRNA	linear	EST 01-OCT-1998
LOCUS	AI076479				
DEFINITION	oz28b04.x1 Soares total_fetus_nb2HF8_9w Homo sapiens cDNA clone IMAGE:1676623 3', similar to contains L1.b1 L1 repetitive element ;				

```

mRNA sequence.
ACCESSION AI076479
VERSION AI076479.1 GI:3405657
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 464)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1450 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 444.
FEATURES
Location/Qualifiers
1..464
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1676623"
/dev_stage="8-9 weeks"
/lab_host="DH108"
/clone_lib="Soares total fetus Nb2HF8 9w"
/note="Vector: pT73-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCCGCTTAATTTTCTTTTCTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "
```

ORIGIN

```

Query Match 5.2%; Score 223; DB 9; Length 464;
Best Local Similarity 100.0%; Pred. No. 7e-95;
Matches 223; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3579 GGACATGTTTCATCAATGCTCCACGAATTTCAACCTCAACATCACTGGGCTGCCAGTTT 3638
DB 464 GGACATGTTTCATCAATGCTCCAGANTTTCACTCAACATCACTGGGCTGCCAGTTT 405
QY 3639 CTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAACCAACATTAAGGA 3698
DB 404 CTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAACCAACATTAAGGA 345
QY 3699 GTACAAGATAGTTTCTCTATGAGAGTTGATTTTCGCAACCAACCAATATCACTTT 3758
DB 344 GTACAAGATAGTTTCTCTATGAGAGTTGATTTTCGCAACCAACCAATATCACTTT 285
QY 3759 CTTTGTGTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 3801
DB 284 CTTTGTGTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 242

RESULT 55
BX644466 500 bp mRNA linear EST 04-SEP-2003
LOCUS DKFZp781N2432 r1 781 (synonym: hlcc4) Homo sapiens cDNA clone
DEFINITION DKFZp781N2432 5', mRNA sequence.
ACCESSION BX644466
VERSION BX644466.1 GI:34478799
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 500)
AUTHORS Bloecker, H., Boecker, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
EST (Bloecker, H., Boecker, M., Mewes, H.W., Weil, B., Amid, C., et al.)
Unpublished (2003)
Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-95764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GfP (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp781N2432) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp781N2432"
/dev_stage="adult"
/lab_host="DH108"
/clone_lib="781 (synonym: hlcc4)"
/note="Vector: pSport1_Sfi; Site_1: SfiI; Site_2: SfiI;
cDNA-collection"
ORIGIN
Query Match 4.3%; Score 185; DB 13; Length 500;
Best Local Similarity 99.6%; Pred. No. 1.1e-76;
Matches 235; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4055 TTCCCAAAACCCATTGACATGGAGCCGTGTTTGGCAACAAAGCCGCTCTCTCTGTGT 4114
DB 1 TTCCCAAAACCCATTGACATGGAGCCGTGTTTGGCAACAAAGCCGCTCTCTCTGTGT 60
QY 4115 TTGTGAGGCTCCCTCGAGGCTGGTGGCATCCCTCTCTCGGAGTCAGTCTGTGTG 4174
DB 61 GTGTGAGGCTCCCTCGAGGCTGGTGGCATCCCTCTCTCGGAGTCAGTCTGTGTG 120
QY 4175 TGCCCAAGCCCTCTGTGTGGACATTTCTCAGCAGATCCGATAGTACAGGAGAGTCAG 4234
DB 121 TGCCCAAGCCCTCTGTGTGGACATTTCTCAGCAGATCCGATAGTACAGGAGAGTCAG 180
QY 4235 GAGCCGTGAGAAACCGGAGAGCAGAGCCCTTGCACAGCCTGGACCTGCATCTGA 4290
DB 181 GAGCCGTGAGAAACCGGAGAGCAGAGCCCTTGCACAGCCTGGACCTGCATCTGA 236

RESULT 56
T69367 517 bp mRNA linear EST 23-FEB-1995
LOCUS yc37g02.r1 Stratagene liver (#937224) Homo sapiens cDNA clone
DEFINITION IMAGE:82898 5', mRNA sequence.
ACCESSION T69367
VERSION T69367.1 GI:680515
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 517)
AUTHORS Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chiappelli, B.,
Chissole, S., Dietrich, N., DuBoque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, M.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevasaki, E., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.
and Marra, M.
```



TITLE	Generation and analysis of 280,000 human expressed sequence tags
JOURNAL	Genome Res. 6 (9), 807-828 (1996)
MEDLINE	97044478
PUBMED	8889549
COMMENT	Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu High quality sequence stops: 219 Source: IMAGE Consortium, LLNL This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. Seq primer: M13RP1 High quality sequence stop: 219. Location/Qualifiers 1. .517 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="GDB:499955" /db_xref="taxon:9606" /clone="IMAGE:82898" /sex="male" /dev_stage="49 years old" /lab_host="SOLR cells (kanamycin resistant)" /clone_lib="Stratagene liver (#937224)" /note="Organ: liver; Vector: pBluescript SK; Site: 1: ECORI; Site: 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Hepatectomy from normal male caucasian. Average insert size: 1.1 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGCGACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
FEATURES	Query Match 4.1%; Score 176; DB 14; Length 517; Best Local Similarity 100.0%; Pred. No. 2.3e-72; Matches 176; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
source	QY 2345 CCCTGCCGCGAAATATCTGTGGCAATTCGGTCGGCAATTTGGTTGGAAATCTCATGTTTGA AAA 2404 Db 59 CCCTGCCGCGAAATATCTGTGGCAATTCGGTCGGCAATTTGGTTGGAAATCTCATGTTTGA AAA 118 QY 2405 TTACTACTGCCAAGGAGAATATGACAAATGCTAAATTTCTCTAGGAAACCAATGCC 2464 Db 119 TTACTACTGCCAAGGAGAATATGACAAATGCTAAATTTCTCTAGGAAACCAATGCC 178 QY 2465 TTTTGGCTTCTCTTACACCCAGAGAGGTAGAAATTTGCTCTTAAGCAGCTGCGA 2520 Db 179 TTTTGGCTTCTCTTACACCCAGAGAGGTAGAAATTTGCTCTTAAGCAGCTGCGA 234
RESULT 57	
BF206327	609 bp mRNA linear EST 06-NOV-2000
LOCUS	60186366AF1 NIH_MGC_19 Homo sapiens CDNA clone IMAGE:4100198 5',
DEFINITION	mRNA sequence.
ACCESSION	BF206327
VERSION	BF206327.1 GI:11099913
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	1. (bases 1 to 609)
TITLE	NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: ATCC CDNA Library Preparation: Ling Hong/Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

Smal; Site 2: Smal; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

	Query Match	3.8%;	Score 165;	DB 13;	Length 265;
	Best Local Similarity	100.0%;	Pred. No. 3.7e-67;		
	Matches 165;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1827	CCATGATGTC	CAACAGATTTGGCCATTCAGCAGCTCTTACACACAGCACCATGATGTGTT	1886	
Db	101	CCATGATGTC	CAACAGATTTGGCCATTCAGCAGCTCTTACACACAGCACCATGATGTGTT	160	
QY	1887	CGGTGGTTTCA	ATAGTCTCTCTCTCAGGCACATCCTGGTATTACCTTCGGAAACAGTGTGA	1946	
Db	161	CGGTGGTTTCA	ATAGTCTCTCTCTCAGGCACATCCTGGTATTACCTTCGGAAACAGTGTGA	220	
QY	1947	TGCGCATCGGAGTGA	AGCCGCTGTTTACGACAGGACCTGGTAT	1991	
Db	221	TGCGCATCGGAGTGA	AGCCGCTGTTTACGACAGGACCTGGTAT	265	

LOCUS	BUI58314	980 bp	mRNA	linear	EST 04-SEP-2002
DEFINITION	AGENCOURT_7775169 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:602544				
ACCESSION	BUI58314				
VERSION	BUI58314.1				
KEYWORDS	EST				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 980)  
NH-MGC <http://mgc.nci.nih.gov/>  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: DCTD/DTP/Gazdar  
cDNA Library Preparation: Life Technologies, Inc

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLM13230 Row: a Column: 09  
High quality sequence start: 8  
High quality sequence stop: 146.  
Location/Qualifiers  
1..980  
source:

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6022544"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NH_MGC_68"
/note="Organ: lung; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.8 kb. Library constructed by Life

```

Technologies. "

ORIGIN

Query Match 3.8%; Score 165; DB 13; Length 980;  
Best Local Similarity 99.5%; Pred. No. 4.9e-67;  
Matches 215; Conservative 0; Indels 0; Gaps 0;

2496 AGAATTGTGCTTTAGCAGCTGCGCAATAATGCAGTCATCTCAGAGCATGTCCAGGCTCAC 2555

```

Db 10 AGAATTGTCTTAAGCAGCTGGAAATAATGCAAGTCAATCTCAGAGCATGTCACAGCTCAC 69
QY 2556 CTTAACCCCGGCTGGCGCTTCGGAGATCAATGTCCTACTGCTGCTGGGAAGATAT 2615
Db 70 CTTAACCCCGGCTGGCGCTTCGGAGATCAATGTCCTACTGCTGCTGGGAAGATAT 129
QY 2616 GTCCCCATTTACAAATAGTTTACTTACAGTGGATGCCCTCTGAGCCAGTCATGCTGGATT 2675
Db 130 GTCCCCATTTACAAATAGTTTACTTACAGTGGATGCCCTCTGAGCCAGTCATGCTGGATT 189
QY 2676 CTGTGGAAATTTATCAGAACCCAGTACTCGGGACT 2711
Db 190 CTGTGGAAATTTATCAGAACCCAGTACTCGGGACT 225

```

```

RESULT 61
T92696 355 bp mRNA linear EST 22-MAR-1995
LOCUS
DEFINITION
IMAGE:118881 5', mRNA sequence.

```

```

ACCESSION
VERSION
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
1 (Bases 1 to 355)
Hillier, L., Lennon, G., Becker, M., Bonaudo, M.F., Chiapelli, B.,
Chissoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierrey-Meg, J.,
Trevasakis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.,
and Marra, M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)

```

```

97044478
8889549
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 271
High quality sequence stops: 255 Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 271 Std Error: 0.00
Seq primer: M13RP1
High quality sequence stop: 255.

```

```

FEATURES
Location/Qualifiers
1..355
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:487170"
/db_xref="taxon:9606"
/clone="IMAGE:118881"
/sex="male"
/dev_stage="72 years"
/lab_host="SOLR cells (kanamycin resistant)"
/clone_lib="Stratagene lung (#937210)"

```

```

Note="Organ: lung; Vector: pBluescript SK-; Site: 1:
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:
Oligo dt. normal lung. Average insert size: 1.0 kb;
Uni-ZAP XR Vector; -5, adaptor sequence: 5' GAATTCGGCAGCAG
3' -3, adaptor sequence: 5' CTCAGATTTTTTTTTTTT 3'"

```

```

ORIGIN
Query Match 3.2%; Score 136; DB 14; Length 355;
Best Local Similarity 100.0%; Pred. No. 3e-53;
Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2568 GGTGGCCCTTCGGAAGATCAATGTCCTACTGCTGCTGGGAAGATATGTCCTATTTAC 2627
Db 73 GGTGGCCCTTCGGAAGATCAATGTCCTACTGCTGCTGGGAAGATATGTCCTATTTAC 132
QY 2628 AAATAGTTTACTTACAGTGGATGCCCTCTGAGCCAGTCATGCTGGATTTCTGGAAATTT 2687
Db 133 AAATAGTTTACTTACAGTGGATGCCCTCTGAGCCAGTCATGCTGGATTTCTGGAAATTT 192
QY 2688 ATCAGAACCCAGTACT 2703
Db 193 ATCAGAACCCAGTACT 208

```

```

RESULT 62
BF475948/c 413 bp mRNA linear EST 05-DEC-2000
LOCUS
DEFINITION
IMAGE:3395955 3', similar to TR:060295 KIAA0548 PROTEIN
; contains LI.D1 LI repetitive element ;, mRNA sequence.
ACCESSION
VERSION
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE
1 (Bases 1 to 413)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

```

```

COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabps@mail.nih.gov
Tissue Procurement: Dr. James Lupski
cDNA Library Preparation: Lupski Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
info@image.llnl.gov
Seq primer: -40UP from Gibco
High quality sequence stop: 364.

```

```

FEATURES
Location/Qualifiers
1..413
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3395955"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/clone_lib="Lupski sciatic nerve"

```

```

Note="Vector: pCMV-SPORT6 (Life Technologies); Site 1:
NotI; Site 2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGCTCCG-3' and
5'-GACTACTTCTAGATCCGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."

```

```

ORIGIN
Query Match 3.1%; Score 135; DB 10; Length 413;
Best Local Similarity 100.0%; Pred. No. 9.4e-53;
Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3667 ATGCTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGAGAAG 3726

```

Db 377 ATGCGCTGTGTTTCAAAACCAACATTAGGAGTACAAAGATAGTTTCTCTAATGAGAG 318  
 QY 3727 TTTGATTTTCGCAACACCAACCAATATCATCTTTCTTTGTTTATGTCAGTAATTCACCTGG 3786  
 Db 317 TTTGATTTTCGCAACACCAACCAATATCATCTTTCTTTGTTTATGTCAGTAATTCACCTGG 258  
 QY 3787 CCCATCAAAATTCAG 3801  
 Db 257 CCCATCAAAATTCAG 243

RESULT 63  
 AA297214  
 LOCUS  
 DEFINITION EST112757 Embryo, 12 week I Homo sapiens cDNA 5' end, mRNA

ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

AA297214 279 bp mRNA linear EST 18-APR-1997  
 EST112757 Embryo, 12 week I Homo sapiens cDNA 5' end, mRNA  
 sequence.  
 AA297214  
 AA297214.1 GI:1949548  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS  
 Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C., Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D., Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S., Glodek,A., Gnehm,C.B., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr., Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M., Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligriano,S.M., Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R., Small,K.V., Spriggs,T.A., Utterback,T.R., Weigman,J.F., Li,Y., Bednarek,D.P., Gao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J., Dinko,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A., He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K., Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H., Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M., Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C., Fraser,C.M., and Venter,J.C.  
 Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence  
 Nature 377 (6547 Suppl), 3-174 (1995)  
 96026280  
 7566098

COMMENT  
 Other ESTs: THCI68226  
 Contact: Kerlavage, AR  
 Bioinformatics  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850 USA  
 Tel: 3018699056  
 Fax: 3018699423  
 Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi.html>)  
 Seq primer: M3 Reverse.  
 Location/Qualifiers  
 1. 279  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="ATCC (inhost):129808"  
 /db\_xref="taxon:9606"  
 /dev\_stage="embryo, 12 wks"  
 /clone\_lib="Embryo, 12 week 1"  
 /note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2: XhoI"

FEATURES  
 source

Query Match  
 Best Local Similarity 100.0%; Pred. No. 6.4e-50;  
 Matches 129; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

QY 2979 TAGTCATTGCTTGGAGCAACCAAGCGTGTGGCTGGTGTACTGATCCAGCAATCTGGCAA 3038  
 Db 42 TAGTCATTGCTTGGAGCAACCAAGCGTGTGGCTGGTGTACTGATCCAGCAATCTGGCAA 101  
 QY 3039 AGGGAATGCATAGAGGGTTCTCTATAAGGACACAGTGAAGATGCTTCGCAAGCCCTTAC 3098  
 Db 102 AGGGAATGCATAGAGGGTTCTCTATAAGGACACAGTGAAGATGCTTCGCAAGCCCTTAC 161  
 QY 3099 AGGAATTT 3107  
 Db 162 AGGAATTT 170

RESULT 64  
 AA771958/c  
 LOCUS  
 DEFINITION

AA771958 370 bp mRNA linear EST 31-DEC-1998  
 ai37e06.s1 Soares parathyroid tumor NBHPA Homo sapiens cDNA clone  
 1359202 3' similar to contains L1.t3 L1 repetitive element ; mRNA  
 sequence.  
 AA771958  
 AA771958.1 GI:2823741  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-re@mail.nih.gov  
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
 Insert Length: 895 Std Error: 0.00  
 Seq primer: -40m3 fwd. Ef from Amersham  
 High quality sequence stop: 297.  
 Location/Qualifiers  
 1. 370  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="1359202"  
 /tissue\_type="parathyroid tumor"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares parathyroid tumor NBHPA"  
 /note="Organ: parathyroid gland; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer  
 [5'-TGTTACCAATCTGAAGTGGGCGCGCCACCAATTTTTTTTTTTTTTTT  
 TTTT-3'] , double-stranded cDNA was size selected, ligated  
 to Eco RI adapters (Pharmacia), digested with Not I and  
 cloned into the Not I and Eco RI sites of a modified pT7T3  
 vector (Pharmacia). Library went through one round of  
 normalization to a Cot = 5. Library constructed by Bento  
 Soares and M.Fatima Bonaldo. RNA from sporadic parathyroid  
 adenomas was kindly provided by Dr. Stephen Marx, National  
 Institute of Diabetes and Digestive and Kidney Diseases,  
 NIH."

FEATURES  
 source

Query Match  
 Best Local Similarity 100.0%; Pred. No. 1.9e-48;  
 Matches 126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

Query Match  
 Best Local Similarity 100.0%; Pred. No. 1.9e-48;  
 Matches 126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

```

3676 GTTTCACCAACCAACATTAAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTT 3735
370 GTTTCACCAACCAACATTAAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTT 311
3736 CGCAACCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTTCACTGGGCCCAATCAAN 3795
310 CGCAACCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTTCACTGGGCCCAATCAAA 251
3796 ATTTCAG 3801
250 ATTTCAG 245

RESULT 65
AA298732 319 bp mRNA linear EST 18-APR-1997
LOCUS EST114356 Supt cells Homo sapiens cDNA 5' end similar to similar to
DEFINITION laminin, B1 chain, mRNA sequence.
ACCESSION AA298732
VERSION AA298732.1 GI:1951085
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 319)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Shall,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Peng,D.-F., Perrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
PUBMED 7566098
COMMENT Other ESTs: THC168226
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 319
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):184690"
/db_xref="taxon:9606"
/cell_type="T-lymphocyte"
/clone_lib="Supt cells"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

FEATURES
source
```

```

Query Match 2.9%; Score 125; DB 9; Length 319;
Best Local Similarity 98.9%; Pred. No. 5.5e-48; Indels 0; Gaps 0;
Matches 275; Conservative 0; Mismatches 3;
QY 3016 ACTGATCCCGAGCAATACTGGCAAGGGAATGTCATAGAGGGTTCCTATATAAGGACCAAGTG 3075
DB 1 ACTGATCCCGAGCAATACTGGCAAGGGAATGTCATAGAGGGTTCCTATATAAGGACCAAGTG 60
QY 3076 AAGATGCTTCGCAAGCCCTACAGGAATTTCTATCCAGAGCCCTGCTCAATTCAGC 3135
DB 61 AAGATGCTTCGCAAGCCCTACAGGAATTTCTATCCAGAGCCCTGCTCAATTCAGC 120
QY 3136 ATGTGTCTAGAGGACAGACATCAACTGTCTTTTCATTCCTGCTTCAGTTGCAATGC 3195
DB 121 ATGTGTCTAGAGGACAGACATCAACTGTCTTTTCATTCCTGCTTCAGTTGCAATGC 180
QY 3196 AACGGCCACAGTAATGTCATCAATCAGAGCAATCTGTGAGAAGTGTGACACCTGACACCA 3255
DB 181 AACGGCCACAGTAATGTCATCAATCAGAGCAATCTGTGAGAAGTGTGACACCTGACACCA 240
QY 3256 GCGAAGCACTCGGAGACCTGCATATCTGGCTTCTACGG 3293
DB 241 GCGAAGCACTCGGAGACCTGCATATCTGGCTTCTACGG 278

RESULT 66
AA297697 230 bp mRNA linear EST 18-APR-1997
LOCUS EST113254 Fetal heart II Homo sapiens cDNA 5' end, mRNA sequence.
DEFINITION AA297697
ACCESSION AA297697.1 GI:1950031
VERSION AA297697.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 230)
Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
Shall,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
Bednarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
Dimke,D., Peng,D.-F., Perrie,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Fraser,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
JOURNAL Nature 377 (6547 Suppl), 3-174 (1995)
MEDLINE 96026280
PUBMED 7566098
COMMENT Other ESTs: THC168226
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 230
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):184690"
/db_xref="taxon:9606"
/cell_type="T-lymphocyte"
/clone_lib="Supt cells"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

FEATURES
source
```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):149184"
/db_xref="taxon:9606"
/dev_stage="fetus"
/clone_lib="Fetal heart II"
/notes="Organ: heart; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

ORIGIN
Query Match      2.9%; Score 123; DB 9; Length 230;
Best Local Similarity 100.0%; Pred. No. 4.6e-47;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2750 TCTGTGAAGGCTGCAACACACAGTGTAAAGCAGTGCAGACACCATGTCCTTGAAGGA 2809
Db 14 TCTGTGAAGGCTGCAACACACAGTGTAAAGCAGTGCAGACACCATGTCCTTGAAGGA 73

QY 2810 CAGCATGTGGAGATTGCACACAGCGCAGCTCTGAGTGCATGTGTGAGCAACATGAAGC 2869
Db 74 CAGCATGTGGAGATTGCACACAGCGCAGCTCTGAGTGCATGTGTGAGCAACATGAAGC 133

QY 2870 AGT 2872
Db 134 AGT 136

RESULT 67
BF700035 602127513F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4284387 5',
LOCUS mRNA sequence.
DEFINITION BF700035
ACCESSION BF700035.1 GI:11985443
VERSION EST.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgc.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM117 row: j column: 04
High quality sequence stop: 615.
FEATURES
Location/Qualifiers
1..683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4284387"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_56"
/notes="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgctcgcc); Site 2: SfiI (ggccattggcc);
Double-stranded cDNA was prepared from cell line RNA.
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTATGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGCGGCGGCACATG-dT(30)BN-3'
(where B = A, C, or G and N = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech

```

```

Laboratories (Palo Alto, CA)."

ORIGIN
Query Match      2.9%; Score 123; DB 10; Length 683;
Best Local Similarity 100.0%; Pred. No. 5.9e-47;
Matches 133; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2836 AGCTCTGAGTGCATGTGTGCAGCAACATGAAGCAGTGTGTGAGCTCCAATGCCCTATGTG 2895
Db 202 AGCTCTGAGTGCATGTGTGCAGCAACATGAAGCAGTGTGTGAGCTCCAATGCCCTATGTG 261

QY 2896 GCTCTCTCCCTTTGGCCAGTGTATGNAATGTTACATGAGCAGCTGCCCTGAA 2955
Db 262 GCTCTCTCCCTTTGGCCAGTGTATGNAATGTTACATGAGCAGCTGCCCTGAA 321

QY 2956 AAT 2958
Db 322 AAT 324

RESULT 68
BZ610572 837 bp DNA linear GSS 08-JUN-2003
LOCUS WHAAD22TR Human MCF7 breast cancer cell line library (MCF7.1) Homo
DEFINITION sapiens genomic clone MCF7_1-1D20, genomic survey sequence.
ACCESSION BZ610572
VERSION BZ610572.1 GI:31519133
KEYWORDS GSS.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 837)
JOURNAL Volik, S., Zhao, S., Chin, K., Brebner, J.H., Herndon, D.R., Tao, Q.,
MEDLINE Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P.,
PUBMED Gray, J.W. and Collins, C.
End-sequence profiling: Sequence-based analysis of aberrant genomes
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
22709111
12788976
COMMENT Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu
This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.
FEATURES
Location/Qualifiers
1..837
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-1D20"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library
(MCF7.1)"
/notes="Vector: pECBAC1; Site_1: HindIII; This library was
constructed from MCF7 breast cancer cell line by Amplicon
Express (http://www.genomex.com) using their standard
procedure."

ORIGIN
Query Match      2.8%; Score 119; DB 28; Length 837;
Best Local Similarity 100.0%; Pred. No. 5e-45;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4049 AGACTGTTCACAAACCATTCGAGCGGTGTTTGCAACAAAGCGCTGTCTCT 4108
Db 36 AGACTGTTCACAAACCATTCGAGCGGTGTTTGCAACAAAGCGCTGTCTCT 95

```

```

QY 4109 CTGTGTTTGTGAGCTCCCTCGAGGCTGGTGGCATCCCTCTCTCTGGGCACTCAGGT 4167
Db 96 CTGTGTTTGTGAGCTCCCTCGAGGCTGGTGGCATCCCTCTCTCTGGGCACTCAGGT 154

RESULT 69
BG945444/c
LOCUS BG945444 134 bp mRNA linear EST 11-JUN-2001
DEFINITION P22-AN0093-190401-008-h10 AN0093 Homo sapiens cDNA, mRNA sequence.
ACCESSION BG945444
VERSION BG945444.1 GI:14344815
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 134)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.P., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F.F., Brentani, R.R., Reis, J.F., de Souza, S.J. and
Simpson, A.J.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.J.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM2&tl2=PM2-AN0093-
190401-008-h10&tl3=2001-04-19&tl4=1)
Seq primer: puc 18 forward
High quality sequence start: 33
High quality sequence stop: 134.
Location/Qualifiers
1..134
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="AN0093"
/note="Organ: amnion normal; Vector: puc18; Site_1: Smal;
Site 2: Smal; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 2.7%; Score 115; DB 12; Length 134;
Best Local Similarity 100.0%; Pred. No. 2.8e-43;
Matches 115; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2606 GGAAGATATGTCGCCATTTACAAATAGTTTACTACAGTGGATCGCTGAGCCCAAGTG 2665
Db 134 GGAAGATATGTCGCCATTTACAAATAGTTTACTACAGTGGATCGCTGAGCCCAAGTG 75

QY 2666 ATGCTGGATTCTGTGGATTTTATCAGAACCCAGTACTCGGGGACTGAAGCTGC 2720
Db 74 ATGCTGGATTCTGTGGATTTTATCAGAACCCAGTACTCGGGGACTGAAGCTGC 20

RESULT 70

```

```

BX669805 530 bp mRNA linear EST 24-OCT-2003
EX669805 Sus Scrofa library (scac) Sus scrofa cDNA clone
scac0029i.b.13 3prim, mRNA sequence.
BX669805
VERSION BX669805.1 GI:37979704
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
REFERENCE 1 (bases 1 to 530)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
Bonnet, A., Tosser-Klopp, G., Benne, F., Cabau, C., Villegier, S.,
Soares, M., Bonaldi, F. and Hately, F.
TITLE A Pig Normalised Multi-Tissue cDNA Library
JOURNAL Unpublished (2003)
COMMENT Contact: Tosser-Klopp G
Genetique Animale
Institut National de la Recherche Agronomique
Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan
cedex, FRANCE
Tel: 33 (0) 5.61.28.51.14
Fax: 33 (0) 5.61.28.53.08
Email: tosser@toulouse.inra.fr
Clone distribution: AGENAE Resource centre, Francois PIUMI,
Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du
Genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex,
FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73
Sequence cleaned of vector, adaptor and repetitions. Contact us
at signenasupport@jouy.inra.fr to obtain the chromatogram of this
sequence.
Plate: 0029 row: b column: 13.
Location/Qualifiers
1..530
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/clone="scac0029i.b.13"
/tissue_type="mixed"
/clone_lib="Sus Scrofa library (scac)"
/note="Vector: p773D-pac vector; tissues: adipose tissue,
brain, kidney, liver, muscle, ovary, testis, heart,
hypothalamus, pancreas, skin, spleen, thymus, placenta,
pituitary gland, seminal vesicle, small intestine,
uterus, adrenals, bulbo uretral gland, cerebral trunk,
epididymis, female gonad, gall-bladder, hippocampus,
large intestine, male gonad, melanocytes, stomach, udder"

ORIGIN
Query Match 2.6%; Score 113; DB 13; Length 530;
Best Local Similarity 100.0%; Pred. No. 3.4e-42;
Matches 113; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 478 TGSGCTCATTGAAGGACGACCAATAGATPAATGAGACTTCGTTTCATCATTTGTGACA 537
Db 387 TGSGCTCATTGAAGGACGACCAATAGATPAATGAGACTTCGTTTCATCATTTGTGACA 328

QY 538 GAGTGTAGTGGGACCAATATATGTTTATGATGGGACTCAATTTATGCACC 590
Db 327 GAGTGTAGTGGGACCAATATATGTTTATGATGGGACTCAATTTATGCACC 275

RESULT 71
CE218575/c
LOCUS CE218575 363 bp DNA linear GSS 25-SEP-2003
DEFINITION tigr-gss-dog-17000326860416 Dog Library Canis familiaris genomic,
genomic survey sequence.
ACCESSION CE218575
VERSION CE218575.1 GI:35374244
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 363)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

REFERENCE  
AUTHORS  
Mammalia; Euthera; Carnivora; Fissipedia; Canidae; Canis.  
1 (bases 1 to 363)  
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,  
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
Venter, J.C.  
The dog genome: survey sequencing and comparative analysis  
Science 301 (5641), 1898-1903 (2003)  
22875432  
PUBMED  
14512627  
COMMENT  
Contact: Kirkness EF  
The Institute for Genomic Research  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-0200  
Fax: 301-838-0208  
Email: ekirknes@tigr.org  
Class: shotgun.

FEATURES  
source  
1.363  
Location/Qualifiers  
/organism="Canis familiaris"  
/mol\_type="genomic DNA"  
/strain="Standard Poodle"  
/db\_xref="taxon:9615"  
/clone\_lib="Dog Library"  
/note="Site 1: BstXI; Libraries were prepared from  
peripheral blood"

ORIGIN  
Query Match 2.6%; Score 110; DB 29; Length 363;  
Best Local Similarity 100.0%; Pred. No. 8.5e-41;  
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2242 CCCAGGATACCCCATGCTACTACTGTACACAGAGACCGAGCTGCGCCCTG 2301  
Db 197 CCCAGGATACCCCATGCTACTACTGTACACAGAGACCGAGCTGCGCCCTG 138  
QY 2302 GACCAGAACTCCAGTGGAGCCCGGAATCAGGAGTGCAATGCGCCGCC 2351  
Db 137 GACCAGAACTCCAGTGGAGCCCGGAATCAGGAGTGCAATGCGCCGCC 88

RESULT 72  
BF316716  
LOCUS  
DEFINITION  
BF316716 942 bp mRNA linear EST 21-NOV-2000  
Accession  
BF316716.1 GI:11265073  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 942)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE  
National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL  
Unpublished (1999)  
COMMENT  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at: [image.llnl.gov](http://image.llnl.gov)  
Plate: L1CM1040 row: j column: 24  
High quality sequence stop: 610.  
Location/Qualifiers  
1.942  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4135799"

FEATURES  
source  
1.942  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4135799"

/tissue type="neuroblastoma"  
/lab host="DH103 (phage-resistant)"  
/clone lib="NIH MGC 19"  
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:  
ECORI; cDNA made by oligo-dT priming. Directionally  
cloned into ECORI/XhoI sites using the following 5'  
adaptor: GGCACAG(G). Library constructed by Ling Hong  
in the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

ORIGIN  
Query Match 2.5%; Score 106; DB 10; Length 942;  
Best Local Similarity 99.4%; Pred. No. 8.7e-39;  
Matches 156; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1238 CAACCTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATATGAGTCATGGGTG 1297  
Db 57 CAACCTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATATGAGTCATGGGTG 116  
QY 1298 TGTTCACCCCTAAGCAAGGAGCAGTATGCACTGGTGGCAGCTCTGCACACATTTGTA 1357  
Db 117 TGTTCACCCCTAAGCAAGGAGCAGTATGCACTGGTGGCAGCTCTGCACACATTTGTA 176  
QY 1358 CACTGAGAAATGCGCGAGTGTGTCATGCTGGTCATCTT 1394  
Db 177 CACTGAGAAATGCGCGAGTGTGTCATGCTGGTCATCTT 213

RESULT 73  
BQ307981  
LOCUS  
DEFINITION  
BQ307981 496 bp mRNA linear EST 16-MAY-2002  
Accession  
BQ307981.1 GI:20846359  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 496)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P.,  
Goldman, G.H., Carvalho, A.P., Matsukuma, A., Baia, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
MEDLINE  
10737800  
PUBMED  
COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR0&t2=MR0-BT5501-250601-001-a09&t3=2001-06-25&t4=1>)  
Seq primer: puc 18 forward  
High quality sequence start: 37  
High quality sequence stop: 115.  
Location/Qualifiers  
1.496  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

FEATURES  
source  
1.496  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"



```
/dev_stage="Adult"
/clone_lib="BT5501"
/note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

ORIGIN
Query Match 2.4%; Score 102; DB 13; Length 496;
Best Local Similarity 100.0%; Pred. No. 6.2e-37; Indels 0; Gaps 0;
Matches 102; Conservative 0; Mismatches 0;

Qy 3091 GCCCTACAGGAATTTCTATCCACGCCCTGCTCAATTCACGATGTCTAGAGGAC 3150
Db 216 GCCCTACAGGAATTTCTATCCACGCCCTGCTCAATTCACGATGTCTAGAGGAC 275

Qy 3151 AGCAGATACACTGCTTTTCAATTCACGTCGCCAGTTCGCCAA 3192
Db 276 AGCAGATACACTGCTTTTCAATTCACGTCGCCAGTTCGCCAA 317

RESULT 74
LOCUS BE171376 160 bp mRNA linear EST 21-JUN-2000
DEFINITION RCI-HT0545-030500-013-d04 HT0545 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE171376
VERSION BE171376.1 GI:8634102
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 160)
Nagai.M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=et2=RCI-HT0545-030
500-013-d04&t3=2000-05-03&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 160.
Location/Qualifiers
1..160
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0545"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
```

```
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN
Query Match 2.3%; Score 97; DB 10; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.2e-34; Indels 0; Gaps 0;
Matches 97; Conservative 0; Mismatches 0;

Qy 4036 GGGGGAGTATAAGACTGTTCCTCCAAACCCATTGCACTGGAGCCGTTTGGCACAAA 4095
Db 126 GGGGGAGTATAAGACTGTTCCTCCAAACCCATTGCACTGGAGCCGTTTGGCACAAA 67

Qy 4096 GCCGCTGCTCTCTCTGTGTGTGTGTGAGGCTCCCTCGAG 4132
Db 66 GCCGCTGCTCTCTCTGTGTGTGTGAGGCTCCCTCGAG 30

RESULT 75
LOCUS AQ826412 517 bp DNA linear GSS 27-AUG-1999
DEFINITION HS 5014 B2 B03 T7A RPCI-11 Human Male BAC Library Homo sapiens
genomic-clone Plate-590 Col=6 Row=D, genomic survey sequence.
ACCESSION AQ826412
VERSION AQ826412.1 GI:5792474
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 517)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieterdejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Research Genetics (info@resgen.com). BAC end web Server:
http://www.htsc.washington.edu
Plate: 590 row: D column: 6
Seq primer: T7
Class: BAC ends
High quality sequence stop: 517.
Location/Qualifiers
1..517
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="plate=590 Col=6 Row=D"
/sex="male"
/clone_lib="RPCI-11 Human Male BAC Library"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"

ORIGIN
Query Match 2.2%; Score 94; DB 28; Length 517;
Best Local Similarity 100.0%; Pred. No. 4.2e-33; Indels 0; Gaps 0;
Matches 94; Conservative 0; Mismatches 0;
```

QY 3323 CATGCAAGTGCAATGGGCACCGCTCTCTGTGCAACACCAACACGGGCAAGTCTTCTGCA 3382  
Db |||||  
QY 340 CATGCAAGTGCAATGGGCACCGCTCTCTGTGCAACACCAACACGGGCAAGTCTTCTGCA 281  
Db |||||  
QY 3383 CCACCAAGGGCGTCAAGGGGACGAGTCCAGCT 3416  
Db |||||  
QY 280 CCACCAAGGGCGTCAAGGGGACGAGTCCAGCT 247  
Db |||||

Search completed: March 1, 2004, 20:45:14  
Job time : 9548 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

March 1, 2004, 07:56:42 ; Search time 1463 Seconds  
(without alignments)  
12457.122 Million cell u

Title: US-09-787-097-13  
 Perfect score: 4290  
 Sequence: 1 atggtggccgcagccgcgcg.....agcctggacctgcacatctga 4290

Scoring table: OLIGO\_NUC  
Gapop\_60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size :

Total number of hits satisfying chosen parameters: 6747726

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 100 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	4290	100.0	4290	3	AAz52276	AAz52276 Human mem
2	4290	100.0	8561	5	AAz72659	AAz72659 DNA encod
3	3801	88.6	3819	3	AAz52277	AAz52277 Human sol
4	3775	88.0	4068	3	AAz52275	AAz52275 Human mem
5	3774	88.0	8590	3	AAz91320	AAz91320 Human mah
6	3285	76.6	4072	3	AAz91921	AAz91921 Human mah
7	3045	71.0	3597	3	AAz52274	AAz52274 Human sol
8	3045	71.0	3597	3	AAz72657	AAz72657 DNA encod
9	3045	71.0	3597	9	ADd70997	ADd70997 Human att
10	1837	42.8	2625	3	AAz91932	AAz91932 Human mah
11	1378	32.1	6373	3	AAz91919	AAz91919 Human mah
12	1077	25.1	3490	5	AAz72660	AAz72660 DNA encod
13	428	10.0	3165	5	AAz878781	AAz878781 DNA encod
14	428	10.0	3313	5	AAz76294	AAz76294 DNA encod
15	314	7.3	668	4	AAz502409	AAz502409 Human sec
16	300	7.0	671	5	AAz72658	AAz72658 DNA encod
17	291	6.8	8827	3	AAz91916	AAz91916 Wild type
18	246	5.7	246	6	ABk38928	ABk38928 cDNA enco
19	246	5.7	246	7	ACa11257	ACa11257 Human lun
20	246	5.7	246	7	ACa02443	ACa02443 Lung canc
21	234	5.5	207433	5	ABz72040	ABz72040 Gene 216
22	234	5.5	207433	4	ABx74891	ABx74891 BAC1098L2
23	228	5.3	481	4	AA114686	AA114686 Human bre



QY 1081 TTCAACCACTCAGATTATAAATGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140  
DB 1081 TTCAACCACTCAGATTATAAATGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140  
QY 1141 CCACCTAAACCCGTTCTGTGGAACAATCTGTTGTTAGATATGTCATTCTTTGSCATTATAC 1200  
DB 1141 CCACCTAAACCCGTTCTGTGGAACAATCTGTTGTTAGATATGTCATTCTTTGSCATTATAC 1200  
QY 1201 AAGGATAAAATTTACATGTATGAGAGAAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260  
DB 1201 AAGGATAAAATTTACATGTATGAGAGAAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260  
QY 1261 TTGAGAGTTTTTCACTTCAATAAGTCACTGCGTGTGTTGACCCCTAAAGGCAAGAGAG 1320  
DB 1261 TTGAGAGTTTTTCACTTCAATAAGTCACTGCGTGTGTTGACCCCTAAAGGCAAGAGAG 1320  
QY 1321 CAGTATGCAAGTGGTGGGCACTCTCGCACACATTTGTACACTGAAGAATGGCCGAGTGGTC 1380  
DB 1321 CAGTATGCAAGTGGTGGGCACTCTCGCACACATTTGTACACTGAAGAATGGCCGAGTGGTC 1380  
QY 1381 ATGCTGGTCACTCTTGGTCACTGCGCTCTCTATGGAATATTAAGCAATGTGCGAGAAATAT 1440  
DB 1381 ATGCTGGTCACTCTTGGTCACTGCGCTCTCTATGGAATATTAAGCAATGTGCGAGAAATAT 1440  
QY 1441 GATTTGGATAGAACACATGGAATATTAACACCCAGGGTGCCCTTGTGCAAGGGGGT 1500  
DB 1441 GATTTGGATAGAACACATGGAATATTAACACCCAGGGTGCCCTTGTGCAAGGGGGT 1500  
QY 1501 TAGCGCCATAGCAGTGTGTTAGCAATAGGACAGGGCCCTATAGTTTCATGTTGGTGTAC 1560  
DB 1501 TAGCGCCATAGCAGTGTGTTAGCAATAGGACAGGGCCCTATAGTTTCATGTTGGTGTAC 1560  
QY 1561 AAGGCTTTTCAGTGCCAAATAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGGAT 1620  
DB 1561 AAGGCTTTTCAGTGCCAAATAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGGAT 1620  
QY 1621 ACCCAGATGTGGACCAATCTTAAGGACAGCGGATTTTCCGTTACTTGCACACAGCTGTG 1680  
DB 1621 ACCCAGATGTGGACCAATCTTAAGGACAGCGGATTTTCCGTTACTTGCACACAGCTGTG 1680  
QY 1681 ATAGTGAATGNAACATGCTGTGTTGGGGGAAACACACAAATGACATCTATGAGC 1740  
DB 1681 ATAGTGAATGNAACATGCTGTGTTGGGGGAAACACACAAATGACATCTATGAGC 1740  
QY 1741 CATGGCGCAATGCTTCTCTCAGATTTCAATGCGCTATGACATTTGCCCTGTGACCGCTGG 1800  
DB 1741 CATGGCGCAATGCTTCTCTCAGATTTCAATGCGCTATGACATTTGCCCTGTGACCGCTGG 1800  
QY 1801 TCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTGACAGTC 1860  
DB 1801 TCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTGACAGTC 1860  
QY 1861 TTACACAACAGCACCATGATGTGTTGCGTGGTTCATATGCTCTCCTCAGCGACATC 1920  
DB 1861 TTACACAACAGCACCATGATGTGTTGCGTGGTTCATATGCTCTCCTCAGCGACATC 1920  
QY 1921 CTGGTATTCACTCGGAACAGTGTGATGCGATCGGATGAAGCGCTGTTGTTAGCAGCA 1980  
DB 1921 CTGGTATTCACTCGGAACAGTGTGATGCGATCGGATGAAGCGCTGTTGTTAGCAGCA 1980  
QY 1981 GGACCTGGTATTCGGTGTGTGGAAACACAGGGTGTCTCAGTGTATCTCGTGGCGCTG 2040  
DB 1981 GGACCTGGTATTCGGTGTGTGGAAACACAGGGTGTCTCAGTGTATCTCGTGGCGCTG 2040  
QY 2041 GCAACTGATGAACAAGAAAGTTAAATCAGAAATGTTTTCAAAAGAACTCTTTGAC 2100  
DB 2041 GCAACTGATGAACAAGAAAGTTAAATCAGAAATGTTTTCAAAAGAACTCTTTGAC 2100  
QY 2101 CATGACAGATGTGACGACACACAGATTTGTATGAGCTGTACAGCCAAACCAATGACTGC 2160  
DB 2101 CATGACAGATGTGACGACACACAGATTTGTATGAGCTGTACAGCCAAACCAATGACTGC 2160  
QY 2161 CACTGGTGAATGACCAATGTTGTGCCAGGAACACAGCTGCTCAGAGGGCCAGATCTCC 2220

DB 2161 CACTGGTGAATGACCAATGTTGTCCCAGGAACACAGCTGCTCAGAAGGCCAGATCTCC 2220  
QY 2221 ATTTTATAGGTATGAGAAATTTGCCCAAGGATAACCCCATGTACTACTGTATAAAGAGACC 2280  
DB 2221 ATTTTATAGGTATGAGAAATTTGCCCAAGGATAACCCCATGTACTACTGTATAAAGAGACC 2280  
QY 2281 AGCTGACGAGTGTGCTCGGACCCAGAACTGCGAGTGGAGCCCGGAATCAGGAGTGC 2340  
DB 2281 AGCTGACGAGTGTGCTCGGACCCAGAACTGCGAGTGGAGCCCGGAATCAGGAGTGC 2340  
QY 2341 ATTGCCCCCGGAAAAATATCTGTGCAATTTGGTGTGCAATTTGTTGAAACTCATGTTTG 2400  
DB 2341 ATTGCCCCCGGAAAAATATCTGTGCAATTTGGTGTGCAATTTGTTGAAACTCATGTTTG 2400  
QY 2401 AAAATTAATACTGCTCAAGGAGAAATATGACAAATGCTTAAATTTGTTCTGTAGGAACCAAT 2460  
DB 2401 AAAATTAATACTGCTCAAGGAGAAATATGACAAATGCTTAAATTTGTTCTGTAGGAACCAAT 2460  
QY 2461 GCCCTTTTGGCTTCTCTTACAAACCCAGAGAGGTAAGTTTCTCCTTAAGCAGCTGCGA 2520  
DB 2461 GCCCTTTTGGCTTCTCTTACAAACCCAGAGAGGTAAGTTTCTCCTTAAGCAGCTGCGA 2520  
QY 2521 ATAATGCAGTCACTCAGAGCATGTCCAACTCACTTAAACCCCATGGGTGCGCTTCGG 2580  
DB 2521 ATAATGCAGTCACTCAGAGCATGTCCAACTCACTTAAACCCCATGGGTGCGCTTCGG 2580  
QY 2581 AAGATCAATGTGCTTACTGCTGCTGGAAAGATATGTCCTTAACTTAACTTACTA 2640  
DB 2581 AAGATCAATGTGCTTACTGCTGCTGGAAAGATATGTCCTTAACTTAACTTACTA 2640  
QY 2641 CAGTGAATGCGCTGCTGAGCCAGTGTGCTGGAATTTCTGGAATTTTATCAGAACCCAGT 2700  
DB 2641 CAGTGAATGCGCTGCTGAGCCAGTGTGCTGGAATTTCTGGAATTTTATCAGAACCCAGT 2700  
QY 2701 ACTCGGGACTGAAGGCTGCACTGCACTCAACCCACTCAATGGTGTGCTGTGAAAGG 2760  
DB 2701 ACTCGGGACTGAAGGCTGCACTGCACTCAACCCACTCAATGGTGTGCTGTGAAAGG 2760  
QY 2761 CCTGCAAAACACAGTGTAGCAGTGTGCGACACCATGTGCTTGAAGGACAGCATGTGGA 2820  
DB 2761 CCTGCAAAACACAGTGTAGCAGTGTGCGACACCATGTGCTTGAAGGACAGCATGTGGA 2820  
QY 2821 GATTGACCAAGCGGAGCTCTGAGTGTGCTGTCAGCAACATGAAGCAGTGTGTTGGAC 2880  
DB 2821 GATTGACCAAGCGGAGCTCTGAGTGTGCTGTCAGCAACATGAAGCAGTGTGTTGGAC 2880  
QY 2881 TCCATGCTTATGTGGCTTCTTCCCTTTTGGCAGTGTATGSAATGTTATACAGTACG 2940  
DB 2881 TCCATGCTTATGTGGCTTCTTCCCTTTTGGCAGTGTATGSAATGTTATACAGTACG 2940  
QY 2941 ACCTGCCCCCTGAAAATTTGTTGAGGCTACTGTACCTGTAGTCAATTTGCTTGAAGCAACCA 3000  
DB 2941 ACCTGCCCCCTGAAAATTTGTTGAGGCTACTGTACCTGTAGTCAATTTGCTTGAAGCAACCA 3000  
QY 3001 GGCTGTGGTGTGTTACTGATCCAGCAATATGCGCAAGGGAATGCAATAGAGGTTCC 3060  
DB 3001 GGCTGTGGTGTGTTACTGATCCAGCAATATGCGCAAGGGAATGCAATAGAGGTTCC 3060  
QY 3061 TATAAGGACCACTGAAGTGTCTTCCAGCCCTTACAGGAATTTCTATCCACAGCCC 3120  
DB 3061 TATAAGGACCACTGAAGTGTCTTCCAGCCCTTACAGGAATTTCTATCCACAGCCC 3120  
QY 3121 CTGCTCAATTTCCAGCATGTGTTAGAGGACAGCAGATACAACTGGTCTTTTCACTCACTGT 3180  
DB 3121 CTGCTCAATTTCCAGCATGTGTTAGAGGACAGCAGATACAACTGGTCTTTTCACTCACTGT 3180  
QY 3181 CGAGCTTGCATGCAACCGCAGCTGAATGATCAATCAGAGCATCTGTGGAAGTGT 3240  
DB 3181 CGAGCTTGCATGCAACCGCAGCTGAATGATCAATCAGAGCATCTGTGGAAGTGT 3240  
QY 3241 GAGAACTGACCAACAGCAAGCACTGCGAGACCTGCAATCTGGCTTCTACCGTGTATCC 3300





QY 2341 ATTGCCCTGCCGAAAAATATCTGTGGCATTTGGCTGGCAATTTGGTTGGAAATCTCATGTTTG 2400  
DB |||||  
QY 2341 ATTGCCCTGCCGAAAAATATCTGTGGCATTTGGCTGGCAATTTGGTTGGAAATCTCATGTTTG 2400  
DB |||||  
QY 2401 AAAATTACTCTGCGAAGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACACAAAT 2460  
DB |||||  
QY 2401 AAAATTACTCTGCGAAGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACACAAAT 2460  
DB |||||  
QY 2461 GCCCTTTTGGCTTCTCTTACAACCCAGAGAGGTAGAAATTTGCTTAAAGCAGCTGGCA 2520  
DB |||||  
QY 2461 GCCCTTTTGGCTTCTCTTACAACCCAGAGAGGTAGAAATTTGCTTAAAGCAGCTGGCA 2520  
DB |||||  
QY 2521 ATAAATGCAATCTCAAGAGCATGTCAGAGTCACTTAAACCCATGCGCTGGCTTGG 2580  
DB |||||  
QY 2521 ATAAATGCAATCTCAAGAGCATGTCAGAGTCACTTAAACCCATGCGCTGGCTTGG 2580  
DB |||||  
QY 2581 AAGATCAATGTCCTTACTGTGCTGCGAAGATATGTCCTCAATTTACAAATAGTTTACTA 2640  
DB |||||  
QY 2581 AAGATCAATGTCCTTACTGTGCTGCGAAGATATGTCCTCAATTTACAAATAGTTTACTA 2640  
DB |||||  
QY 2641 CAGTGATGCGCTCTGAGCCAGTGTGCTGATTTCTGTGAATTTTATCAAAACCCAGT 2700  
DB |||||  
QY 2641 CAGTGATGCGCTCTGAGCCAGTGTGCTGATTTCTGTGAATTTTATCAAAACCCAGT 2700  
DB |||||  
QY 2701 ACTCGGGGACTGAAGGCTGCAACCTGCAATCAACCCACTCAATGAGTGTCTGTGAAAG 2760  
DB |||||  
QY 2701 ACTCGGGGACTGAAGGCTGCAACCTGCAATCAACCCACTCAATGAGTGTCTGTGAAAG 2760  
DB |||||  
QY 2761 CTTGCAAAACCAAGTCTAAGCAGTCCGCGACACCACTGTGCTTGAGGACAGCATGTGGA 2820  
DB |||||  
QY 2761 CTTGCAAAACCAAGTCTAAGCAGTCCGCGACACCACTGTGCTTGAGGACAGCATGTGGA 2820  
DB |||||  
QY 2821 GATTGACACGCGGAGCTCTGAGTGCAATGCTGCTGAGTGTGCTGAGCAACATGAAGCAGTGTGAGC 2880  
DB |||||  
QY 2821 GATTGACACGCGGAGCTCTGAGTGCAATGCTGCTGAGTGTGCTGAGCAACATGAAGCAGTGTGAGC 2880  
DB |||||  
QY 2881 TCCAAATGCTATGTGCGCTCTTCTTCCCTTTGCGCAGTGTGAGTGTGATACAGATGAGC 2940  
DB |||||  
QY 2881 TCCAAATGCTATGTGCGCTCTTCTTCCCTTTGCGCAGTGTGAGTGTGATACAGATGAGC 2940  
DB |||||  
QY 2941 ACCTGCCCCCTGAAAAATTTGTTGAGGCTACTGTACCTGTAGTCACTTGTGAGCAACCA 3000  
DB |||||  
QY 2941 ACCTGCCCCCTGAAAAATTTGTTGAGGCTACTGTACCTGTAGTCACTTGTGAGCAACCA 3000  
DB |||||  
QY 3001 GGCTGTGGCTGTGCTGATCCAGCAATCTGCGAAGGCAATGATAGAGGTTCC 3060  
DB |||||  
QY 3001 GGCTGTGGCTGTGCTGATCCAGCAATCTGCGAAGGCAATGATAGAGGTTCC 3060  
DB |||||  
QY 3061 TATAAAGGACCAAGTGAAGTGTCTTCCGAGCCCTACAGGAAATTTCTATCCAGGCC 3120  
DB |||||  
QY 3061 TATAAAGGACCAAGTGAAGTGTCTTCCGAGCCCTACAGGAAATTTCTATCCAGGCC 3120  
DB |||||  
QY 3121 CTGCTCAATTCAGCATGTGTAGAGGACAGCAGATACAACTGGTCTTTTCACTTCACTGT 3180  
DB |||||  
QY 3121 CTGCTCAATTCAGCATGTGTAGAGGACAGCAGATACAACTGGTCTTTTCACTTCACTGT 3180  
DB |||||  
QY 3181 CCAGCTTGCCAAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT 3240  
DB |||||  
QY 3181 CCAGCTTGCCAAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGT 3240  
DB |||||  
QY 3241 GAGAACCTTGACCAAGGCAAGCAGTGGAGACCTGCAATCTTGGCTTCTAGCGTATCCC 3300  
DB |||||  
QY 3241 GAGAACCTTGACCAAGGCAAGCAGTGGAGACCTGCAATCTTGGCTTCTAGCGTATCCC 3300  
DB |||||  
QY 3301 ACCAATGAGGGAATGTGACGATGCAAGTGTGCAATGCGCTCTCTGTGCAACACC 3360  
DB |||||  
QY 3301 ACCAATGAGGGAATGTGACGATGCAAGTGTGCAATGCGCTCTCTGTGTGCAACACC 3360  
DB |||||  
QY 3361 AACACGGGCAAGTGTCTTGACCAACCAAGGCGGTCAAGGGGGACGAGTGCAGCTATGT 3420  
DB |||||  
QY 3361 AACACGGGCAAGTGTCTTGACCAACCAAGGCGGTCAAGGGGGACGAGTGCAGCTATGT 3420  
DB |||||

QY 3421 GAGGTAGAAAAATCGATACCAAGAAACCTCTCAGAGGAACATGTTATTTATATCTCTTCTT 3480  
DB |||||  
QY 3421 GAGGTAGAAAAATCGATACCAAGAAACCTCTCAGAGGAACATGTTATTTATATCTCTTCTT 3480  
DB |||||  
QY 3481 ATTGACTATCAGTTCACTTTAGTCTATCCAGGAGATGATCGTATTATACAGCTATC 3540  
DB |||||  
QY 3481 ATTGACTATCAGTTCACTTTAGTCTATCCAGGAGATGATCGTATTATACAGCTATC 3540  
DB |||||  
QY 3541 AATTTTGTGGCTACTCTCTGACGAAACAAAACAGGGATTTGGACATGTTCTCAATGCTCTC 3600  
DB |||||  
QY 3541 AATTTTGTGGCTACTCTCTGACGAAACAAAACAGGGATTTGGACATGTTCTCAATGCTCTC 3600  
DB |||||  
QY 3601 AAGAAATTTCAACCTCAACATCACTGGGCTGCGAGTTCTCAGCTGGAAACCCAGGCTGGA 3660  
DB |||||  
QY 3601 AAGAAATTTCAACCTCAACATCACTGGGCTGCGAGTTCTCAGCTGGAAACCCAGGCTGGA 3660  
DB |||||  
QY 3661 GAAGAGATGCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720  
DB |||||  
QY 3661 GAAGAGATGCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3720  
DB |||||  
QY 3721 GAGAAATTTGATTTTGGCAACCAACCAATATACATTTCTTTGTTATGTCAGTAATTC 3780  
DB |||||  
QY 3721 GAGAAATTTGATTTTGGCAACCAACCAATATACATTTCTTTGTTATGTCAGTAATTC 3780  
DB |||||  
QY 3781 ACCTGGCCCATCAAAATTCAGATTTGCTTCTCTCAGCAGCAAAATTTATGACCTGGTA 3840  
DB |||||  
QY 3781 ACCTGGCCCATCAAAATTCAGATTTGCTTCTCTCAGCAGCAAAATTTATGACCTGGTA 3840  
DB |||||  
QY 3841 CAGTTCTCTGTAATTTTCAAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3900  
DB |||||  
QY 3841 CAGTTCTCTGTAATTTTCAAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3900  
DB |||||  
QY 3901 TGGAAATCAAAACAAAGTTGTTGGGCTCCAGACGTAGAGAGCACTTCTTCGAGAGATG 3960  
DB |||||  
QY 3901 TGGAAATCAAAACAAAGTTGTTGGGCTCCAGACGTAGAGAGCACTTCTTCGAGAGATG 3960  
DB |||||  
QY 3961 CAACAGATGCGCAGCGCTCCCTTTGCTCTGTAATGTCGCTTGGAAACAGATGAGGAG 4020  
DB |||||  
QY 3961 CAACAGATGCGCAGCGCTCCCTTTGCTCTGTAATGTCGCTTGGAAACAGATGAGGAG 4020  
DB |||||  
QY 4021 CTTCTCTGATCTTATTGGGGGAGTATAAGACTGTTCCCAACCCATTCATTCAGTGGCGC 4080  
DB |||||  
QY 4021 CTTCTCTGATCTTATTGGGGGAGTATAAGACTGTTCCCAACCCATTCATTCAGTGGCGC 4080  
DB |||||  
QY 4081 TGTTTTGGCAACAAAGCCGCTGCTCTCTCTGTTTGTGAGGCTCCCTCGAGGCTGGGT 4140  
DB |||||  
QY 4081 TGTTTTGGCAACAAAGCCGCTGCTCTCTCTGTTTGTGAGGCTCCCTCGAGGCTGGGT 4140  
DB |||||  
QY 4141 GGCATCCCTCTCTCTGGGAGTCAAGTCTTGTGTGGCCAGCGCTCTGGTGGACATTTCT 4200  
DB |||||  
QY 4141 GGCATCCCTCTCTCTGGGAGTCAAGTCTTGTGTGGCCAGCGCTCTGGTGGACATTTCT 4200  
DB |||||  
QY 4201 CAGCAGATCCGATAGTGTACAGGAGAGTCAAGAGCGTCAAGAAACCGGAAGCAGCAG 4260  
DB |||||  
QY 4201 CAGCAGATCCGATAGTGTACAGGAGAGTCAAGAGCGTCAAGAAACCGGAAGCAGCAG 4260  
DB |||||  
QY 4261 CCCCCCTGCAAGCCTGGGACCTGCAATCTGA 4290  
DB |||||  
QY 4261 CCCCCCTGCAAGCCTGGGACCTGCAATCTGA 4290  
DB |||||

## RESULT 3

AAZ52277

ID AAZ52277 standard; cDNA; 3819 BP.

XX AAZ52277;  
XX ACXX 18-JUL-2000 (first entry)  
XX Human soluble attractin-2 cDNA.XX Human; soluble attractin-2; immune response; macrophage; monocyte;  
XX T cell; immunosuppressed patient; cancer;

KW





1441 GATTTGGATAAGAACACATATGGAGTATATATACACACCCAGGGTGCCTTTGTCAAGGGGGT 1500  
1441 GATTTGGATAAGAACACATATGGAGTATATATACACACCCAGGGTGCCTTTGTCAAGGGGGT 1500  
1501 TAGGGCCATAGCAGTGTATACACCATAGGACCCAGGGCCCTATACGTTTCATGGTGGCTAC 1560  
1501 TAGGGCCATAGCAGTGTATACACCATAGGACCCAGGGCCCTATACGTTTCATGGTGGCTAC 1560  
1561 AAGGCTTTTCAGTCCCAATAAGTACCGGCTTGACAGATGATCTCTACCGATATGATGTGGAT 1620  
1561 AAGGCTTTTCAGTCCCAATAAGTACCGGCTTGACAGATGATCTCTACCGATATGATGTGGAT 1620  
1621 ACCAGATGTGACCAATCTTTAAGACAGCCGATTTTTCGGTATCTTGACACAGAGCTGTG 1680  
1621 ACCAGATGTGACCAATCTTTAAGACAGCCGATTTTTCGGTATCTTGACACAGAGCTGTG 1680  
1681 ATAGTGAGTGGAAACCATGCTGGTGTGGGGGAAACACACACATGACATCTATGAGC 1740  
1681 ATAGTGAGTGGAAACCATGCTGGTGTGGGGGAAACACACACATGACATCTATGAGC 1740  
1741 CATGGCCCAATGCTTCTTCATGATTTCAATGCTATGCTTCAATGCTATGCTTCAATGCTATGCT 1800  
1741 CATGGCCCAATGCTTCTTCATGATTTCAATGCTATGCTTCAATGCTATGCTTCAATGCTATGCT 1800  
1801 TCAGTGTCTCCAGACCATGATCTCCACCATGATGTCACAGATTTGGCCATTCAGCAGTC 1860  
1801 TCAGTGTCTCCAGACCATGATCTCCACCATGATGTCACAGATTTGGCCATTCAGCAGTC 1860  
1861 TTACACACAGCACCACATGATGTGTGGTGGTGTTCATAGTCTCTCCCTCAGCAGATC 1920  
1861 TTACACACAGCACCACATGATGTGTGGTGGTGTTCATAGTCTCTCCCTCAGCAGATC 1920  
1921 CTGGTATTTCACTCGGACAGTGTATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1980  
1921 CTGGTATTTCACTCGGACAGTGTATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1980  
1981 GGACCTGTATTCGGTGTGTGGAAACACAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2040  
1981 GGACCTGTATTCGGTGTGTGGAAACACAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2040  
2041 GCACTGTATGAACAGAGAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGAC 2100  
2041 GCACTGTATGAACAGAGAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGAC 2100  
2101 CATGACAGATGTGACAGACACACAGATTTGTACAGCTGTACAGCAACACCAATGACTGC 2160  
2101 CATGACAGATGTGACAGACACACAGATTTGTACAGCTGTACAGCAACACCAATGACTGC 2160  
2161 CACTGGTGCATGACATTTGTGTCCTCCAGGAAACACAGCTGCTCAGAAAGCCAGATCTCC 2220  
2161 CACTGGTGCATGACATTTGTGTCCTCCAGGAAACACAGCTGCTCAGAAAGCCAGATCTCC 2220  
2221 ATTTTGTAGTATGAATTTGCTCCAGAGTAAACCCATGACTACTGTACAGAGAGACC 2280  
2221 ATTTTGTAGTATGAATTTGCTCCAGAGTAAACCCATGACTACTGTACAGAGAGACC 2280  
2281 AGCTGCAGAGCTGTGCTCCCTGGACAGAACTGCCAGTGGAGCCCGGAAATCAGGAGTGC 2340  
2281 AGCTGCAGAGCTGTGCTCCCTGGACAGAACTGCCAGTGGAGCCCGGAAATCAGGAGTGC 2340  
2341 ATTGCGCTCCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTGGAAACTCATGTTTG 2400  
2341 ATTGCGCTCCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTGGAAACTCATGTTTG 2400  
2401 AAAATTTACTATGCGCAAGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAAT 2460  
2401 AAAATTTACTATGCGCAAGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAAT 2460  
2461 GCCCTTTGGCTTCTCTTAACCCAGAGAGAGTAAATTTGCTCTTAAGAGAGCTGGCA 2520  
2461 GCCCTTTGGCTTCTCTTAACCCAGAGAGAGTAAATTTGCTCTTAAGAGAGCTGGCA 2520

2521 ATAAATCAGTCTATCTCAGAGCATGTCCAAAGCTACCTTAAACCCCATGGTCCGCTTCGG 2580  
2521 ATAAATCAGTCTATCTCAGAGCATGTCCAAAGCTACCTTAAACCCCATGGTCCGCTTCGG 2580  
2581 AGATCAATGTGTCTCTACTGGTGGGAGATATGTCCCAATTTACAAATAGTTACTA 2640  
2581 AGATCAATGTGTCTCTACTGGTGGGAGATATGTCCCAATTTACAAATAGTTACTA 2640  
2641 CAGTGTATCCGCTCTGAGCCCAAGTGTGATGCTGTGGAATTTTATCAGAACCCAGT 2700  
2641 CAGTGTATCCGCTCTGAGCCCAAGTGTGATGCTGTGGAATTTTATCAGAACCCAGT 2700  
2701 ACTCGGGGACTGAGGCTGCAACCTGCACTCAACCCACTCAATGGTAGTGTCTGTGAAGG 2760  
2701 ACTCGGGGACTGAGGCTGCAACCTGCACTCAACCCACTCAATGGTAGTGTCTGTGAAGG 2760  
2761 CCTGCAAAACACAGTCTTAAGCAGTCCCGACACCATGTGCTTGGAGCAGCATGTGGA 2820  
2761 CCTGCAAAACACAGTCTTAAGCAGTCCCGACACCATGTGCTTGGAGCAGCATGTGGA 2820  
2821 GATTCACACAGCGGACAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGGAC 2880  
2821 GATTCACACAGCGGACAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGGAC 2880  
2881 TCCAAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGTTATACAGTATGAC 2940  
2881 TCCAAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGTTATACAGTATGAC 2940  
2941 ACCTGCCCCCTGAAATTTGTTTCCAGCTACTGTACTCTGTAGTCAATTTGTTGGAGCAACA 3000  
2941 ACCTGCCCCCTGAAATTTGTTTCCAGCTACTGTACTCTGTAGTCAATTTGTTGGAGCAACA 3000  
3001 GGCTGTGGTGTGTACTGATCCAGCAATCTGCAAGGGAATGCAATGAGGGTTCC 3060  
3001 GGCTGTGGTGTGTACTGATCCAGCAATCTGCAAGGGAATGCAATGAGGGTTCC 3060  
3061 TATAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGGCC 3120  
3061 TATAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGGCC 3120  
3121 CTGCTCAATTCAGCATGTGTCTAGAGCAGCAGATACACTGCTCTTTTCACTCTGCT 3180  
3121 CTGCTCAATTCAGCATGTGTCTAGAGCAGCAGATACACTGCTCTTTTCACTCTGCT 3180  
3181 CCAGCTTGCATGCAATGCAAGGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAAGT 3240  
3181 CCAGCTTGCATGCAATGCAAGGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAAGT 3240  
3241 GAGACCTGACACAGGCAAGCAGTGCAGACCTGCAATCTGCGTCTTACGGTATGCC 3300  
3241 GAGACCTGACACAGGCAAGCAGTGCAGACCTGCAATCTGCGTCTTACGGTATGCC 3300  
3301 ACCAATGAGGGAATGTGAGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 3360  
3301 ACCAATGAGGGAATGTGAGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 3360  
3361 AACACGGGCAAGTGTCTGCAACCAAGGGGCTCAAGGGGAGAGTGCAGCTATGT 3420  
3361 AACACGGGCAAGTGTCTGCAACCAAGGGGCTCAAGGGGAGAGTGCAGCTATGT 3420  
3421 GAGGTAGAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTTATATCTCTCTT 3480  
3421 GAGGTAGAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTTATATCTCTCTT 3480  
3481 ATTGACTATCAGTTCACCTTTAGTCTATCCAGGGAAGTATCGCTATTTACACAGTATC 3540  
3481 ATTGACTATCAGTTCACCTTTAGTCTATCCAGGGAAGTATCGCTATTTACACAGTATC 3540  
3541 AATTTTGGCTTCTCTGACAAACAGGGAATTTGGACATGTTTCATCAATGCCTCC 3600  
3541 AATTTTGGCTTCTCTGACAAACAGGGAATTTGGACATGTTTCATCAATGCCTCC 3600  
3601 AAGAAATTTCAACCTCAACATCAGTGGGCTGCCAGTCTCTCAGCTGGAAACCCAGGCTGGA 3660

Db 3601 AAGAATTCAACCTCAACATCACTGGCTGCCAGTTTCTACGTGGAACCCAGGCTGGA 3660  
Qy 3661 GAAGAGATGCTGTGTTTTCAAAACCAACATTAAAGAGGTACAAAGATAGTTTCTCTAAT 3720  
Db 3661 GAAGAGATGCTGTGTTTTCAAAACCAACATTAAAGAGGTACAAAGATAGTTTCTCTAAT 3720  
Qy 3721 GAGAAGTTGATTTTGCAACCCACCAAAATATCACATTTCTTTGTTTATGTAGTAAATTC 3780  
Db 3721 GAGAAGTTGATTTTGCAACCCACCAAAATATCACATTTCTTTGTTTATGTAGTAAATTC 3780  
Qy 3781 ACCTGGCCCATCAAAATTCAG 3801  
Db 3781 ACCTGGCCCATCAAAATTCAG 3801  
RESULT 4  
AAZ52275 ID AAZ52275 standard; cDNA; 4068 BP.  
XX AC AAZ52275;  
XX 18-JUL-2000 (first entry)  
XX Human membrane attractin-1 cDNA.  
XX Human; membrane attractin-1; immune response; macrophage; monocyte;  
XX T cell; immunostimulant; immunosuppressed patient; cancer;  
XX immunodeficiency syndrome; transplant; autoimmune disease; ss.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
XX CDS 1..4068  
XX /\*tag= a  
XX /product= "Membrane attractin-1"  
XX WO200015651-A1.  
XX 23-MAR-2000.  
XX 14-SEP-1999; 99WO-US020948.  
XX 14-SEP-1998; 98US-0100137P.  
XX (DAND ) DANA FARBER CANCER INST INC.  
XX Duke-Cohan JS, Schlossman SF;  
XX WPI; 2000-271373/23.  
XX P-PSDB; RAY70690.  
XX Isolated nucleic acids encoding human attractin polypeptides useful for  
XX enhancing immune responses.  
XX Claim 3; Fig 10; 120pp; English.  
XX The patent discloses four forms of human attractin polypeptides which  
XX enhance immune response by promoting macrophage and monocyte spreading in  
XX the presence of T cells. These include soluble attractin-1 and -2 and  
XX membrane attractin-1 and -2. These various forms of attractin are encoded  
XX by alternatively spliced mRNA molecule transcribed from a single gene.  
XX The present sequence is a cDNA encoding human membrane attractin-1.  
XX Membrane attractin differs from soluble attractin in having a  
XX transmembrane domain and a cytoplasmic domain at the C-terminal. This  
XX sequence can be used to enhance immune response in immunosuppressed  
XX patients such as those undergoing chemo- and radio-therapy treatment for  
XX cancer or those suffering from common variable immunodeficiency syndrome.  
XX The proteins may also be used to screen modulators (agonists and  
XX antagonists) of immune response which may also be used to regulate immune  
XX reactions. Attractin antibodies can be used to inhibit immune response in  
XX transplant recipients or patients afflicted with autoimmune disease

SQ Sequence 4068 BP; 1042 A; 948 C; 1027 G; 1051 T; 0 U; 0 Other;  
Query Match 88.0%; Score 3775; DB 3; Length 4068;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3975; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
Qy 312 CTGTGTCACACGGCGTCCCTGCACACCCCTGGCACCAGGCGCTCCGCCCGCGGCTG 371  
Db 90 CTGTGTCACACGGCGTCCCTGCACACCCCTGGCACCAGGCGCTCCGCCCGCGGCTG 149  
Qy 372 GGTGGGCGAGCAATGCCAGCACTGCGGGGGCGCTTCAGACTAACTGGATCTTCTGGGTT 431  
Db 150 GGTGGGCGAGCAATGCCAGCACTGCGGGGGCGCTTCAGACTAACTGGATCTTCTGGGTT 209  
Qy 432 TGTGACAGATGACCTCGAAATTAATAACAAACAGAGTGCACGTGGCTCATTAAGG 491  
Db 210 TGTGACAGATGACCTCGAAATTAATAACAAACAGAGTGCACGTGGCTCATTAAGG 269  
Qy 492 ACAGCAAAATAGATAATGAGACTTCGTTTCAATCAATTTGTACAGAGTGTAGTTGGGA 551  
Db 270 ACAGCAAAATAGATAATGAGACTTCGTTTCAATCAATTTGTACAGAGTGTAGTTGGGA 329  
Qy 552 CCATTTATATGTTTATGATGGGACTCAATTTATGACACCGCTAGTTGCTGATTTAGTGG 611  
Db 330 CCATTTATATGTTTATGATGGGACTCAATTTATGACACCGCTAGTTGCTGATTTAGTGG 389  
Qy 612 CCTCATTTGCTGAGAGAGATGGCAATGAGACTGCTCCCTGAGGTTGTTGCCACATCAGG 671  
Db 390 CCTCATTTGCTGAGAGAGATGGCAATGAGACTGCTCCCTGAGGTTGTTGCCACATCAGG 449  
Qy 672 TTATGCTTGTGTCATTTTTTTTAGTGATGCTCTTATAATTTGACTGGATTTAATTTAC 731  
Db 450 TTATGCTTGTGTCATTTTTTTTAGTGATGCTCTTATAATTTGACTGGATTTAATTTAC 509  
Qy 732 TTACAGTTTGTATGTTGTCGCAATTAATGCTCAGGCGAGGAGAGTGTAGATCAGTAA 791  
Db 510 TTACAGTTTGTATGTTGTCGCAATTAATGCTCAGGCGAGGAGAGTGTAGATCAGTAA 569  
Qy 792 TAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAACTGGAAGCTGAAGCATGTGACAT 851  
Db 570 TAGCAGCGATACTGTTGAATGTGAATGTTCTGAAACTGGAAGCTGAAGCATGTGACAT 629  
Qy 852 TCCTCATCTGACAGCAACTGTGGTTTCTCATCGAGGCATCTGCAATTCAGTGTGT 911  
Db 630 TCCTCATCTGACAGCAACTGTGGTTTCTCATCGAGGCATCTGCAATTCAGTGTGT 689  
Qy 912 CAGAGGATGCTCCTGCTCTCAGACTGCGAGGTCCTGGATGTTGATGTTCTGTACCAGC 971  
Db 690 CAGAGGATGCTCCTGCTCTCAGACTGCGAGGTCCTGGATGTTGATGTTCTGTACCAGC 749  
Qy 972 TAACCACTCATTTTGGACTCGAGAGCAATTTCTAACTTAAAGCTCCCGAGAGCATCTCA 1031  
Db 750 TAACCACTCATTTTGGACTCGAGAGCAATTTCTAACTTAAAGCTCCCGAGAGCATCTCA 809  
Qy 1032 TAAAGCTGTGTCAATGGAAACATTTATGGGTTGTTGGAGGATATATGTTCAACCACTC 1091  
Db 810 TAAAGCTGTGTCAATGGAAACATTTATGGGTTGTTGGAGGATATATGTTCAACCACTC 869  
Qy 1092 AGATTATACATGTTCTAGCGTATGACCTTCTTCTAGGGAGTGGCTTCCACTAAACCG 1151  
Db 870 AGATTATACATGTTCTAGCGTATGACCTTCTTCTAGGGAGTGGCTTCCACTAAACCG 929  
Qy 1152 TTCTGTGAACAATGTGGTTGTAGATATGTTGATGTTCTTTGGCATTTATCAAGGATAAAT 1211  
Db 930 TTCTGTGAACAATGTGGTTGTAGATATGTTGATGTTCTTTGGCATTTATCAAGGATAAAT 989  
Qy 1212 TTACATGTATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGTAGTGTAGAGTTT 1271  
Db 990 TTACATGTATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGTAGTGTAGAGTTT 1049  
Qy 1272 TCACATTCATAATGAGTCAATGGTGTGTGTACCCCTAAGGCAAGGAGCATGTAGTCACT 1331  
Db 1050 TCACATTCATAATGAGTCAATGGTGTGTGTACCCCTAAGGCAAGGAGCATGTAGTCACT 1109

QY 1332 GGTGGGCACTCTGCACACATGTTTACACTGAAGATGGCCGAGTGGTCAATGCTGGTCAAT 1391  
DB 1110 GGTGGGCACTCTGCACACATGTTTACACTGAAGATGGCCGAGTGGTCAATGCTGGTCAAT 1169  
QY 1392 CTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAA 1451  
DB 1170 CTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAA 1229  
QY 1452 GAACATGAGAGTATATACACACCCAGGGTCCCTTGTGCAAGGGGTTACGGCCATAG 1511  
DB 1230 GAACATGAGAGTATATACACACCCAGGGTCCCTTGTGCAAGGGGTTACGGCCATAG 1289  
QY 1512 CAGTGTATACGACCATAGGACCGGCCCTATAGTTCATGGTGGCTACAGGCTTTCAG 1571  
DB 1290 CAGTGTATACGACCATAGGACCGGCCCTATAGTTCATGGTGGCTACAGGCTTTCAG 1349  
QY 1572 TGCCAAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTG 1631  
DB 1350 TGCCAAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTG 1409  
QY 1632 GACCATTTTAAGGACACGCCAATTTTCCGTTACTTGCACACAGCTGTGATGATGAGTGG 1691  
DB 1410 GACCATTTTAAGGACACGCCAATTTTCCGTTACTTGCACACAGCTGTGATGATGAGTGG 1469  
QY 1692 AACCATGCTGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCAA 1751  
DB 1470 AACCATGCTGTGTTTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCAA 1529  
QY 1752 ATGCTTCTCTTACAGATTTTATGGCTATGACATTCCTGTGACCGCTGGTCAATGCTTCC 1811  
DB 1530 ATGCTTCTCTTACAGATTTTATGGCTATGACATTCCTGTGACCGCTGGTCAATGCTTCC 1589  
QY 1812 CAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGCAGCTTTACACAACAG 1871  
DB 1590 CAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGCAGCTTTACACAACAG 1649  
QY 1872 CACCATGATGTTTGGTGGTGTCAATAGTCTCTCTCAGGACATCTGATTTAC 1931  
DB 1650 CACCATGATGTTTGGTGGTGTCAATAGTCTCTCTCAGGACATCTGATTTAC 1709  
QY 1932 CTCGGAACAGTGCATCGGATCGGAGTGAAGCGCTTGTAGCAGCAGGACTGGTAT 1991  
DB 1710 CTCGGAACAGTGCATCGGATCGGAGTGAAGCGCTTGTAGCAGCAGGACTGGTAT 1769  
QY 1992 TCGGTGTGTGGAACACAGGCTCGTCTCAGTGTATCTCGTGGCGCTGCAACTGATGA 2051  
DB 1770 TCGGTGTGTGGAACACAGGCTCGTCTCAGTGTATCTCGTGGCGCTGCAACTGATGA 1829  
QY 2052 ACAGAGAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTTTCAGCAGTGCAGATG 2111  
DB 1830 ACAGAGAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTTTCAGCAGTGCAGATG 1889  
QY 2112 TGACCAACACACAGATTTTACAGCTGTACAGCCACACCAATGACTGGCCACTGGTGCAA 2171  
DB 1890 TGACCAACACACAGATTTTACAGCTGTACAGCCACACCAATGACTGGCCACTGGTGCAA 1949  
QY 2172 TGACCAATGTGTCCCAAGGAACACACAGCTGTACAGAGGCCAGATCTCCATTTTATAGTA 2231  
DB 1950 TGACCAATGTGTCCCAAGGAACACACAGCTGTACAGAGGCCAGATCTCCATTTTATAGTA 2009  
QY 2232 TGAGAAATGCCCCAAGGATACCCCATGACTACTGTAAACAGAACAGCAGCTGCAAGGAG 2291  
DB 2010 TGAGAAATGCCCCAAGGATACCCCATGACTACTGTAAACAGAACAGCAGCTGCAAGGAG 2069  
QY 2292 CTGTGCCCTGGAACAGAACTGCCAGTGGGAGCCCCCGAATCAGGAGTGCATTGCCCTGCC 2351  
DB 2070 CTGTGCCCTGGAACAGAACTGCCAGTGGGAGCCCCCGAATCAGGAGTGCATTGCCCTGCC 2129  
QY 2352 CGAAATATCTGGGCAATGGCTGGCAATTTGGTTGGAACCTCATGTTGAAATTTACTAC 2411  
DB 2130 CGAAATATCTGGGCAATGGCTGGCAATTTGGTTGGAACCTCATGTTGAAATTTACTAC 2189

QY 2412 TGCCAAAGGAGAAATTTATGACAAATGCTAAATTTGTTCTGTAGGAACCAATGCCCTTTTGGC 2471  
DB 2190 TGCCAAAGGAGAAATTTATGACAAATGCTAAATTTGTTCTGTAGGAACCAATGCCCTTTTGGC 2249  
QY 2472 TTCTCTTACAAACCCAGAGAGGTAGAAATTTGTTCTTAAGCAGCTGCGAAATAATGCAGTC 2531  
DB 2250 TTCTCTTACAAACCCAGAGAGGTAGAAATTTGTTCTTAAGCAGCTGCGAAATAATGCAGTC 2309  
QY 2532 ATCTCAGAGCATGTCCAAAGCTCACCTTAAACCCATGGGTGGCCCTTCGGAAGATCAATGT 2591  
DB 2310 ATCTCAGAGCATGTCCAAAGCTCACCTTAAACCCATGGGTGGCCCTTCGGAAGATCAATGT 2369  
QY 2592 GTCTACTGTGCTCGGAGATATGTCCTCAATTTTCAAAATAGTTTACTACAGTGGATGCC 2651  
DB 2370 GTCTACTGTGCTCGGAGATATGTCCTCAATTTTCAAAATAGTTTACTACAGTGGATGCC 2429  
QY 2652 GTCTGAGCCCACTGATGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACT 2711  
DB 2430 GTCTGAGCCCACTGATGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACT 2489  
QY 2712 GAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGTGGAAGCCCTGCAAAACA 2771  
DB 2490 GAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGTGGAAGCCCTGCAAAACA 2549  
QY 2772 CAGTGTCTAAGCAGTGCAGGACACCATGTGCTTGAAGCAGCATGTGAGATTTGCACCAG 2831  
DB 2550 CAGTGTCTAAGCAGTGCAGGACACCATGTGCTTGAAGCAGCATGTGAGATTTGCACCAG 2609  
QY 2832 CGGAGCTCTGAGTGCATGTGTCAGCAACAATGAAGCAGTGTGAGTGCATTCATGCTA 2891  
DB 2610 CGGAGCTCTGAGTGCATGTGTCAGCAACAATGAAGCAGTGTGAGTGCATTCATGCTA 2669  
QY 2892 TGTGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACAGTACAGCTGCCCCC 2951  
DB 2670 TGTGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACAGTACAGCTGCCCCC 2729  
QY 2952 TGAAAATTTGTTACAGCTACTGTACTGTAGTCAATGTTGGAGCAACAGGCTGTGGCTG 3011  
DB 2730 TGAAAATTTGTTACAGCTACTGTACTGTAGTCAATGTTGGAGCAACAGGCTGTGGCTG 2789  
QY 3012 GTGTACTCATCCAGCAATCTGTCAGAGGAATGCAATAGAGGTTTCTTAAGGACC 3071  
DB 2790 GTGTACTCATCCAGCAATCTGTCAGAGGAATGCAATAGAGGTTTCTTAAGGACC 2849  
QY 3072 AGTGAAGATGCTTCGAAAGCCCTTACAGGAAATTTCTTATCCAGGCCCTGCTCAATTC 3131  
DB 2850 AGTGAAGATGCTTCGAAAGCCCTTACAGGAAATTTCTTATCCAGGCCCTGCTCAATTC 2909  
QY 3132 CAGCATGTCTTAGAGGACAGATCAACTGGTCTTTCATTTCACTGTCAGCTTGGCA 3191  
DB 2910 CAGCATGTCTTAGAGGACAGATCAACTGGTCTTTCATTTCACTGTCAGCTTGGCA 2969  
QY 3192 ATGCAACCGCCACAGTAATCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGAC 3251  
DB 2970 ATGCAACCGCCACAGTAATCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGAC 3029  
QY 3252 CACAGGAAGCAGTGCAGAGACTGTGATCTGCTTACGCTGATCCCAATGAGG 3311  
DB 3030 CACAGGAAGCAGTGCAGAGACTGTGATCTGCTTACGCTGATCCCAATGAGG 3089  
QY 3312 GAAATGTGAGCCATGCAAGTGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGAC 3371  
DB 3090 GAAATGTGAGCCATGCAAGTGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGAC 3149  
QY 3372 GTGCTTCTGCAACCAAGGGGCTCAAGGGGAGCAGGTGCGAGCTATGTGAGGTAGAAA 3431  
DB 3150 GTGCTTCTGCAACCAAGGGGCTCAAGGGGAGCAGGTGCGAGCTATGTGAGGTAGAAA 3209  
QY 3432 TCGATACCAAGGAACCTCTCAGAGGAACATGTTATATATCTCTTATTTGACTATCA 3491  
DB 3210 TCGATACCAAGGAACCTCTCAGAGGAACATGTTATATATCTCTTATTTGACTATCA 3269  
QY 3492 GTTCACCTTTTGTCTATCCAGGAAGATGATGCTTATACAGCTATCAATTTTGTGGC 3551





QY 2893 GTGGCCCTCCCTCCCTTTGGCCAGTGATGAATGATATACGATGAGCACCTGCCCCCT 2952  
Db |||||  
QY 2890 GTGCCCTCCCTTTCCCTTTGGCCAGTGATGAATGATATACGATGAGCACCTGCCCCCT 2949  
Db |||||  
QY 2953 GAAATTTGTTTCCAGCTACTGTACCTGTAGTCAITTCCTTGGAGCAACACAGGCTGTGGCTGG 3012  
Db |||||  
QY 2950 GAAATTTGTTTCCAGCTACTGTACCTGTAGTCAITTCCTTGGAGCAACACAGGCTGTGGCTGG 3009  
Db |||||  
QY 3013 TGTAATGATCCAGCAATCTGGCAAGGGAATGATGAGGATTCCTATTAAGGACCA 3072  
Db |||||  
QY 3010 TGTAATGATCCAGCAATCTGGCAAGGGAATGATGAGGATTCCTATTAAGGACCA 3069  
Db |||||  
QY 3073 GTGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTC 3132  
Db |||||  
QY 3070 GTGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTC 3129  
Db |||||  
QY 3133 AGCATGTGTAGAGGACAGCAGATACAATCTGGTCTTTCTATCTGTCTCAGCTTGGCAA 3192  
Db |||||  
QY 3130 AGCATGTGTAGAGGACAGCAGATACAATCTGGTCTTTCTATCTGTCTCAGCTTGGCAA 3189  
Db |||||  
QY 3193 TGCAACGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACC 3252  
Db |||||  
QY 3190 TGCAACGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACC 3249  
Db |||||  
QY 3253 ACAGCAAGCACTCCGAGACCTGCATATCTGGCTTCTACGGTATCCCAATGAGGG 3312  
Db |||||  
QY 3250 ACAGCAAGCACTCCGAGACCTGCATATCTGGCTTCTACGGTATCCCAATGAGGG 3309  
Db |||||  
QY 3313 AAATGTAGCCATCGAATGGAAGGCAAGCTCTGTGCAACACCAACAGGGCAAG 3372  
Db |||||  
QY 3310 AAATGTAGCCATCGAATGGAAGGCAAGCTCTGTGCAACACCAACAGGGCAAG 3369  
Db |||||  
QY 3373 TGCTTCTGCACCAACCAAGGGCGTCAAGGGGACAGTGCAGCTATGTGAGGTAGAAAT 3432  
Db |||||  
QY 3370 TGCTTCTGCACCAACCAAGGGCGTCAAGGGGACAGTGCAGCTATGTGAGGTAGAAAT 3429  
Db |||||  
QY 3433 CGATACAGAAACCTCTCAGAGGACATGTTATTAATCTTCTATGATATGATCAG 3492  
Db |||||  
QY 3430 CGATACAGAAACCTCTCAGAGGACATGTTATTAATCTTCTTATGATATCAG 3489  
Db |||||  
QY 3493 TTCACTTTAGTCTATCCAGAGAGATGATCGCTATTACACAGCTATCAATTTGTGGCT 3552  
Db |||||  
QY 3490 TTCACTTTAGTCTATCCAGAGAGATGATCGCTATTACACAGCTATCAATTTGTGGCT 3549  
Db |||||  
QY 3553 ACTCTGACGAAACAAACAGGATTTGGAATGTTATCAATGCTTCAAGAAATTTCAAC 3612  
Db |||||  
QY 3550 ACTCTGACGAAACAAACAGGATTTGGAATGTTATCAATGCTTCAAGAAATTTCAAC 3609  
Db |||||  
QY 3613 CTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGATGCTCT 3672  
Db |||||  
QY 3610 CTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGATGCTCT 3669  
Db |||||  
QY 3673 GTTGTTCGAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAATTTGAT 3732  
Db |||||  
QY 3670 GTTGTTCGAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAATTTGAT 3729  
Db |||||  
QY 3733 TTGCGACACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTTCACTGGCCCATC 3792  
Db |||||  
QY 3730 TTGCGACACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTTCACTGGCCCATC 3789  
Db |||||  
QY 3793 AAATTCAGATTCCTCTCTCAGCACAGCAATTTTATGACCTGGTACAGTTTCTTCGTG 3852  
Db |||||  
QY 3790 AAATTCAGATTCCTCTCTCAGCACAGCAATTTTATGACCTGGTACAGTTTCTTCGTG 3849  
Db |||||  
QY 3853 ACTTCTTCAGTGTGTTTCTCTCTTTGCTCTGGTGTGCTGTGTTTGAAGATCAAA 3912  
Db |||||  
QY 3850 ACTTCTTCAGTGTGTTTCTCTCTTTGCTCTGGTGTGCTGTGTTTGAAGATCAAA 3909  
Db |||||  
QY 3913 CAAGTCTTCGGCTCCAGAGTAGAGACACTTCTTCAGAGATGCAACAGATGGCC 3972  
Db |||||  
QY 3910 CAAGTCTTCGGCTCCAGAGTAGAGACACTTCTTCAGAGATGCAACAGATGGCC 3969  
Db |||||  
QY 3973 AGCGCTCCCTTTGCCCTCTGTAATGTGCGCTTTGGAACAGATGAGGACCTTCCTGATCTT 4032

Db 3970 AGCGCTCCCTTTGCCCTCTGTAATGTGCGCTTTGGAACAGATGAGGACCTTCCTGATCTT 4029  
QY 4033 ATTGGGGGAGTATAAAGACTGTTCCAAAACCCATTGACATGAGACCGCTGTTTGGCAAC 4092  
Db 4030 ATTGGGGGAGTATAAAGACTGTTCCAAAACCCATTGACATGAGACCGCTGTTTGGCAAC 4089  
QY 4093 AAAGCCGCTGTCCTCTCTGTTGTTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCCT 4152  
Db 4090 AAAGCCGCTGTCCTCTCTGTTGTTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCCT 4149  
QY 4153 CTGGGCGAGTCAAGTCTGCTGTGCTGCGCCAGCGCCCTGTTGGAACATTTCTCAGCAGATGCGG 4212  
Db 4150 CTGGGCGAGTCAAGTCTGCTGTGCTGCGCCAGCGCCCTGTTGGAACATTTCTCAGCAGATGCGG 4209  
QY 4213 ATAGTGTACAAGGAGAAGTCAAGAGCCGTGAGAAACCCGAGACAGAGCCCTTCACAG 4272  
Db 4210 ATAGTGTACAAGGAGAAGTCAAGAGCCGTGAGAAACCCGAGACAGAGCCCTTCACAG 4269  
QY 4273 CTGGGACCTGCATCTGA 4290  
Db 4270 CTGGGACCTGCATCTGA 4287

RESULT 6  
AAZ91921  
ID AAZ91921 standard; cDNA; 4072 BP.  
XX  
AC AAZ91921;  
XX  
DT 08-JUN-2000 (first entry)  
XX  
DE Human mahogany protein coding sequence #3.  
XX  
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;  
KW weight regulation; cell therapy; body weight disorder; cachexia;  
KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
KW hyperphagia; Anticobesity; antianorexic; anticachexic; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200005373-A2.  
XX  
PD 03-FEB-2000.  
XX  
PF 21-JUL-1999; 99WO-US016484.  
XX  
PR 21-JUL-1998; 98US-0093630P.  
PR 20-OCT-1998; 98US-0104978P.  
PR 05-FEB-1999; 99US-00245041.  
XX  
PA (MILL-) MILLENIUM PHARM INC.  
XX  
PI Moore K, Nagle DL;  
XX  
DR WPI: 2000-195103/17.  
DR P-PSDB; AAY81808.  
XX  
PT New human and murine mahogany genes, useful, e.g. for diagnosis and  
PT treatment of body weight disorders.  
XX  
PS Claim 1; Fig 19a; 180pp; English.  
XX  
CC This sequence represents a human mahogany gene of the invention. The  
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
CC proteins (ii); (iii) as a source of diagnostic probes and primers for  
CC therapeutics; (iv) as a source of mutations, regulatory defects, in  
CC detecting expression of mg genes or mutations, regulatory defects, in  
CC this gene, or for isolation of related sequences; and (iv) in (cell-  
CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to  
CC identify other (extra)cellular products involved in weight regulation,  
CC and to screen for agents that disrupt interaction between (ii) and other  
CC macromolecules. The Ab are used to detect abnormal levels (or function)



CC of (II) (for diagnosis, prognosis or monitoring of treatment); to  
CC evaluate (II)-expressing cells intended for cell therapy, and as  
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
CC mg polypeptide) are used to identify agents (A) that modulate mg  
CC activity. (A) are potentially useful for the treatment of body weight  
CC disorders, particularly obesity, cachexia or anorexia, or other  
CC conditions associated with the mg gene such as hyperpigmentation,  
CC hyperphagia and disorders that result in increased metabolic rate  
XX  
SQ Sequence 4072 BP; 1087 A; 940 C; 1045 G; 1000 T; 0 U; 0 Other;  
Query Match 76.6%; Score 3285; DB 3; Length 4072;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3485; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 313 TGTGTCAACGGGGTGTGCTGCAACCTGGACCGGCGAGTCGGCTGCGCCGGGTGG 372  
DB 310 TGTGTCAACGGGGTGTGCTGCAACCTGGACCGGCGAGTCGGCTGCGCCGGGTGG 369  
QY 373 GTGGCGAGCAATGCCAGACTGGGGGGCGGCTTCAGACTAACTGGATCTTCTGGTTT 432  
DB 370 GTGGCGAGCAATGCCAGACTGGGGGGCGGCTTCAGACTAACTGGATCTTCTGGTTT 429  
QY 433 GTGACAGATGACCTGGAAATTAATAACAAAGAGTGCAGTGGCTCAATTGAGGA 492  
DB 430 GTGACAGATGACCTGGAAATTAATAACAAAGAGTGCAGTGGCTCAATTGAGGA 489  
QY 493 CAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGAC 552  
DB 490 CAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGAC 549  
QY 553 CATTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCATTTAGTGGC 612  
DB 550 CATTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCATTTAGTGGC 609  
QY 613 CTCATTGTTCTGACAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT 672  
DB 610 CTCATTGTTCTGACAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT 669  
QY 673 TATGCCCTGCTGCATTTTATGATGATGCTGCTCTTATTAATTTGACTGGATTAATTAAT 732  
DB 670 TATGCCCTGCTGCATTTTATGATGATGCTGCTCTTATTAATTTGACTGGATTAATTAAT 729  
QY 733 TACAGTTTGTATGATGCTCCAAATAACTGTCAGCGCGAGGAGTGTAAAGTCAAGTAAT 792  
DB 730 TACAGTTTGTATGATGCTCCAAATAACTGTCAGCGCGAGGAGTGTAAAGTCAAGTAAT 789  
QY 793 AGCAGCGAACTGTTGAAATGTGAATGTTCTGAAACTGGAAAGTGAAGCATGTGACATT 852  
DB 790 AGCAGCGATACTGTTGAAATGTGAATGTTCTGAAACTGGAAAGTGAAGCATGTGACATT 849  
QY 853 CCTCACTGTACAGCAACTGTTTCCCTCATCGAGCATCTGCAATTCAGTGTATGTC 912  
DB 850 CCTCACTGTACAGCAACTGTTTCCCTCATCGAGCATCTGCAATTCAGTGTATGTC 909  
QY 913 AGAGATGCTCTGCTTCTCAGACTGGCAGGCTCTGATGTTTCAAGTTCCTGATCCAGCT 972  
DB 910 AGAGATGCTCTGCTTCTCAGACTGGCAGGCTCTGATGTTTCAAGTTCCTGATCCAGCT 969  
QY 973 AACAGTCAATTTGGACTCGAGAGGAATATTTCACTTAAGCTTCCCGCAGGATCTCAT 1032  
DB 970 AACAGTCAATTTGGACTCGAGAGGAATATTTCACTTAAGCTTCCCGCAGGATCTCAT 1029  
QY 1033 AAAGCTGTGCTCAATGGAACAATTTATGTTGGTGTGTTGAGGATATGTTCAACACTCA 1092  
DB 1030 AAAGCTGTGCTCAATGGAACAATTTATGTTGGTGTGTTGAGGATATGTTCAACACTCA 1089  
QY 1093 GATTATTAACATGTTTCTAGCTATGACCTTCTTAGGGAGTGGCTTCCACTAAACCGT 1152  
DB 1090 GATTATTAACATGTTTCTAGCTATGACCTTCTTAGGGAGTGGCTTCCACTAAACCGT 1149  
QY 1153 TCTGTGAACAATGTGTTGTTAGATATGTTTCACTTCTTGGCATTTATACAGGATTAAT 1212

DB 1150 TCTGTGAACAATGTGTTGTTAGATATGTTCTTCTTGGCAATTATACAAGGATAAAT 1209  
QY 1213 TACATGTATGGAGGAAAAATTCATCCACTGGGAATGTGACCAATGAGTTGAGAGTTTT 1272  
DB 1210 TACATGTATGGAGGAAAAATTCATCCACTGGGAATGTGACCAATGAGTTGAGAGTTTT 1269  
QY 1273 CACATTCAATAAGTATGAGTTCATGGGTGTTGTTGACCCCTAAGGCAAGAGCAGTATCAGTG 1332  
DB 1270 CACATTCAATAAGTATGAGTTCATGGGTGTTGTTGACCCCTAAGGCAAGAGCAGTATCAGTG 1329  
QY 1333 GTTGGCACTCTGCACACATTTTACACTGAAGATGGCGAGTGGTCACTGCTGCTATC 1392  
DB 1330 GTTGGCACTCTGCACACATTTTACACTGAAGATGGCGAGTGGTCACTGCTGCTATC 1389  
QY 1393 TTTGGTCACTGCCCCCTCTCTATGTATATAAGCAATGTGCAGGAATATGATTTGGATAAG 1452  
DB 1390 TTTGGTCACTGCCCCCTCTCTATGTATATAAGCAATGTGCAGGAATATGATTTGGATAAG 1449  
QY 1453 AACATGTAGATATATACACACCCAGGTTGCCCTTGTGCAAGGGGTTCACGCCATAGC 1512  
DB 1450 AACATGTAGATATATACACACCCAGGTTGCCCTTGTGCAAGGGGTTCACGCCATAGC 1509  
QY 1513 AGTGTTTAGACCATAGGACAGGCGCCCTATACGTTTCATGTTGGCTCAAGGCTTTCAGT 1572  
DB 1510 AGTGTTTAGACCATAGGACAGGCGCCCTATACGTTTCATGTTGGCTCAAGGCTTTCAGT 1569  
QY 1573 GCCAATAAGTACCGGTTGCAGATGATCTTACCAATATGATGTGGATACCCAGATGTGG 1632  
DB 1570 GCCAATAAGTACCGGTTGCAGATGATCTTACCAATATGATGTGGATACCCAGATGTGG 1629  
QY 1633 ACCATTCTTAAGACACAGCGATTTTCCGTTTACTTGTGACACAGCTGTGATGAGTGA 1692  
DB 1630 ACCATTCTTAAGACACAGCGATTTTCCGTTTACTTGTGACACAGCTGTGATGAGTGA 1689  
QY 1693 ACCATGCTGGTGTGTTGGGGGAAACACACACATATGATCTATGAGCCATGGCGCCAAA 1752  
DB 1690 ACCATGCTGGTGTGTTGGGGGAAACACACACATATGATCTATGAGCCATGGCGCCAAA 1749  
QY 1753 TGTCTTCTTTCAGATTTTCAATGCTGATGACATTTGCTGTGACCGCTGCTGAGTCTCC 1812  
DB 1750 TGTCTTCTTTCAGATTTTCAATGCTGATGACATTTGCTGTGACCGCTGCTGAGTCTCC 1809  
QY 1813 AGACCTGTATCTCCACCATGATGTCAACAGATTTGCGCATTCAGCAGTCTTACCAACAGC 1872  
DB 1810 AGACCTGTATCTCCACCATGATGTCAACAGATTTGCGCATTCAGCAGTCTTACCAACAGC 1869  
QY 1873 ACCATGTATGTTTCCGTTGTTTCAATAGTCTTCTTCCCTCAGGACATCTCTGGTATTCAC 1932  
DB 1870 ACCATGTATGTTTCCGTTGTTTCAATAGTCTTCTTCCCTCAGGACATCTCTGGTATTCAC 1929  
QY 1933 TCGGAACAGTGTGATGCGCATCGGAGTGAAGCGCTGTTTGTAGCAGCAGGACCTGGTATT 1992  
DB 1930 TCGGAACAGTGTGATGCGCATCGGAGTGAAGCGCTGTTTGTAGCAGCAGGACCTGGTATT 1989  
QY 1993 CGGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGGCTGGCAACTGATGAA 2052  
DB 1990 CGGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGGCTGGCAACTGATGAA 2049  
QY 2053 CAAGAAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCATGACAGATGT 2112  
DB 2050 CAAGAAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCATGACAGATGT 2109  
QY 2113 GACAGCACACAGATTTGTACAGCTGTACAGCAACACCAATGATGCTGCACTGGTCAAT 2172  
DB 2110 GACAGCACACAGATTTGTACAGCTGTACAGCAACACCAATGATGCTGCACTGGTCAAT 2169  
QY 2173 GACATTTGTCCCGAGAAACACAGCTGCTCAGNAGCCAGATCTCCATTTTAGGTAT 2232  
DB 2170 GACCATTTGTCCCGAGAAACACAGCTGCTCAGNAGCCAGATCTCCATTTTAGGTAT 2229  
QY 2233 GAGAAATTCGCCCAAGGATAACCCCTATGCTACTTGTACTGTAAACAAGAGCAGCTGCAGGAGC 2292  
DB 2230 GAGAAATTCGCCCAAGGATAACCCCTATGCTACTTGTACTGTAAACAAGAGCAGCTGCAGGAGC 2289



2293 TGTCCTTGGACAGAACTGCCAGTGGAGCCCGGAAATCAGAGTGCATTGGCCCTGCC 2352  
Db TGTCCTTGGACAGAACTGCCAGTGGAGCCCGGAAATCAGAGTGCATTGGCCCTGCC 2349  
2293 GAAATATCTGTGCACTGGCTGCAATTTGGTGGAACTCATGTTTGGAAATTAATCTACT 2412  
Db GAAATATCTGTGCACTGGCTGCAATTTGGTGGAACTCATGTTTGGAAATTAATCTACT 2409  
2413 GCCAAGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGAACCAATGCCCCCTTTGGCT 2472  
Db GCCAAGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGAACCAATGCCCCCTTTGGCT 2469  
2473 TCTCTTCAACCCAGAGAGAGTGAATTTGCTCTTAAGCAGCTGCGAAATAATGCACTCA 2532  
Db TCTCTTCAACCCAGAGAGAGTGAATTTGCTCTTAAGCAGCTGCGAAATAATGCACTCA 2529  
2533 TCTCAGAGCATGTCCAGCTCACTTAACCCCAATGGTGGCCCTTCGGAAGATCAATGTG 2592  
Db TCTCAGAGCATGTCCAGCTCACTTAACCCCAATGGTGGCCCTTCGGAAGATCAATGTG 2589  
2593 TCCTACTGGTGGAGAGATATGTCCTTCAAAATAGTTTACTACAGTGCATGCCG 2652  
Db TCCTACTGGTGGAGAGATATGTCCTTCAAAATAGTTTACTACAGTGCATGCCG 2649  
2590 TCCTACTGGTGGAGAGATATGTCCTTCAAAATAGTTTACTACAGTGCATGCCG 2649  
2653 TCTCAGCCCACTGATGCTGCAATTTCTGGAATTTTATCAGAACCCAGTACTCGGGACTG 2712  
Db TCTCAGCCCACTGATGCTGCAATTTCTGGAATTTTATCAGAACCCAGTACTCGGGACTG 2709  
2713 AAGGCTGCACTGCATCAACCCCACTCAATGGTGTCTGTGAAAGGCTGCAACCCAC 2772  
Db AAGGCTGCACTGCATCAACCCCACTCAATGGTGTCTGTGAAAGGCTGCAACCCAC 2769  
2773 AGTCTAAGCAGTCCCGACACCACTGTGCTTGAAGCAGCATGTGGAGATTGCACAGC 2832  
Db AGTCTAAGCAGTCCCGACACCACTGTGCTTGAAGCAGCATGTGGAGATTGCACAGC 2829  
2833 GGCAGCTCTGAGTGCATGTGGTGAGCAACATGAAGCAGTGTGGAGTCCCAATGCCAT 2892  
Db GGCAGCTCTGAGTGCATGTGGTGAGCAACATGAAGCAGTGTGGAGTCCCAATGCCAT 2889  
2893 GTGGCCCTCTTCCCTTTGGCAGTGTATGAATGGTATACGATGACACCTGCCCCCT 2952  
Db GTGGCCCTCTTCCCTTTGGCAGTGTATGAATGGTATACGATGACACCTGCCCCCT 2949  
2953 GAAATTTGTTCCAGCTACTGTACCTGTAGTCAATTTGAGCAACAGAGCTGTGGCTGG 3012  
Db GAAATTTGTTCCAGCTACTGTACCTGTAGTCAATTTGAGCAACAGAGCTGTGGCTGG 3009  
3013 TGTAATGATCCAGCAATCTGGCAAGGAAATGCAATGAGGTTCTCTATAAGGACCA 3072  
Db TGTAATGATCCAGCAATCTGGCAAGGAAATGCAATGAGGTTCTCTATAAGGACCA 3069  
3073 GTGAAGATGCTTCCGAGCCCTTACAGGAAATTTCTATCCAGCCCTGTCTCAATTC 3132  
Db GTGAAGATGCTTCCGAGCCCTTACAGGAAATTTCTATCCAGCCCTGTCTCAATTC 3129  
3133 AGCATGTCTAGAGGACAGAGATACAACTGTGTTTCAATCTGTCCAGCTTGCCTCA 3192  
Db AGCATGTCTAGAGGACAGAGATACAACTGTGTTTCAATCTGTCCAGCTTGCCTCA 3189  
3193 TGCAAGCCCACTAGTAATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACC 3252  
Db TGCAAGCCCACTAGTAATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACC 3249  
3253 ACAGCAAGCACTGCGAGACTGATATCTGGCTTCTACGGTGTATCCCAATTGGAGGG 3312  
Db ACAGCAAGCACTGCGAGACTGATATCTGGCTTCTACGGTGTATCCCAATTGGAGGG 3309  
3313 AAATGTGAGCCATGCAATGGGACCGCTCTGTGCAACACACACACAGGGCAAG 3372  
Db AAATGTGAGCCATGCAATGGGACCGCTCTGTGCAACACACACACAGGGCAAG 3369

3373 TGCTTCTGCACCAAGGGCGTCAAGGGGACGAGTGCAGCTATGTGAGTGAATAAT 3432  
Db TGCTTCTGCACCAAGGGCGTCAAGGGGACGAGTGCAGCTATGTGAGTGAATAAT 3429  
3433 CGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTATCTCTTATTGACATCAG 3492  
Db CGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTATCTCTTATTGACATCAG 3489  
3493 TTCACTTTAGTCTATCCAGGAAAGATGATCGTATTACAGCTATCAATTTTGGCT 3552  
Db TTCACTTTAGTCTATCCAGGAAAGATGATCGTATTACAGCTATCAATTTTGGCT 3549  
3553 ACTCTGACGAAACAAACAGGGATTGGACATGTTTCATCAATGCCCTCCAAGAAATTCAC 3612  
Db ACTCTGACGAAACAAACAGGGATTGGACATGTTTCATCAATGCCCTCCAAGAAATTCAC 3609  
3613 CTCACATCAGCTGGCTGCCAGTTTCTCAGCTGAAACCCAGGCTGGAGAGAGATGCT 3672  
Db CTCACATCAGCTGGCTGCCAGTTTCTCAGCTGAAACCCAGGCTGGAGAGAGATGCT 3669  
3673 GTTGTGTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGTAT 3732  
Db GTTGTGTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGTAT 3729  
3733 TTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTCACCTGGCCATC 3792  
Db TTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTCACCTGGCCATC 3789  
3793 AAAATTTCAG 3801  
Db AAAATTTCAG 3798

## RESULT 7

AAZ52274  
ID AAZ52274 standard; cDNA; 3597 BP.

XX AAZ52274;

XX 19-JUL-2000 (first entry)

XX Human soluble attractin-1 cDNA.

XX Human; soluble attractin-1; immune response; macrophage; monocyte;  
T cell; immunosuppressed patient; cancer;  
immunodeficiency syndrome; transplant; autoimmune disease; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..3597

XX /tag= a

XX /product= "Soluble attractin-1"

XX WO200015651-A1.

XX PN 23-MAR-2000.

XX XX 14-SEP-1999; 99WO-US020949.

XX PR 14-SEP-1998; 98US-0100137P.

XX (DAND ) DANA FARBER CANCER INST INC.

XX Duke-Cohan JS, Schlossman SF;

XX WPI; 2000-271373/23.

XX DR P-PSDB; AAY70689.

XX Isolated nucleic acids encoding human attractin polypeptides useful for  
enhancing immune responses.

XX Claim 3; Fig 8; 120pp; English.

XX The patent discloses four forms of human attractin polypeptides which  
CC enhance immune response by promoting macrophage and monocyte spreading in  
CC the presence of T cells. These include soluble attractin-1 and -2 and  
CC membrane attractin-1 and -2. These various forms of attractin are encoded  
CC by alternatively spliced mRNA molecule transcribed from a single gene.  
CC The present sequence is a cDNA encoding soluble attractin-1 obtained from  
CC activated T cell and human fetal liver libraries. This sequence can be  
CC used to enhance immune response in immunosuppressed patients such as  
CC those undergoing chemo- and radio-therapy treatment for cancer or those  
CC suffering from common variable immunodeficiency syndrome. The proteins  
CC may also be used to screen modulators (agonists and antagonists) of  
CC immune responses which may also be used to regulate immune reactions.  
CC Attractin antibodies can be used to inhibit immune response in transplant  
CC recipients or patients afflicted with autoimmune disease  
XX  
SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 71.0%; Score 3045; DB 3; Length 3597;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 3485; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY	312	CTGTGTCAACGGCGTGTCTGCAACCTGGCACCGCCAGTGGCTCTGCCCGCGCGCTG	371
DB	90	CTGTGTCAACGGCGTGTCTGCAACCTGGCACCGCCAGTGGCTCTGCCCGCGCGCTG	149
QY	372	GGTGGGAGCAATGCCAGCACTCGGGGGCGGTTT	431
DB	150	GGTGGGAGCAATGCCAGCACTCGGGGGCGGTTT	209
QY	432	TGTGACAGATGGACTGGAAATTAATACAAACGAAGTGCAGCGGTCAATGAGG	491
DB	210	TGTGACAGATGGACTGGAAATTAATACAAACGAAGTGCAGCGGTCAATGAGG	269
QY	492	ACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGA	551
DB	270	ACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGA	329
QY	552	CCATTTATATTTTATGATGGGACTCAATTTATGACCGGTAGTGTCTGCATTTAGTGG	611
DB	330	CCATTTATATTTTATGATGGGACTCAATTTATGACCGGTAGTGTCTGCATTTAGTGG	389
QY	612	CCTCATTTCTCTGAGAGATGCAATGAGACTGTCCCTGAGTGTGTGCCACATCAGG	671
DB	390	CCTCATTTCTCTGAGAGATGCAATGAGACTGTCCCTGAGTGTGTGCCACATCAGG	449
QY	672	TTATGCCCTTGTGCATTTTATGATGCTGTCTTAAATTTGACTGGATTTAATATTAC	731
DB	450	TTATGCCCTTGTGCATTTTATGATGCTGTCTTAAATTTGACTGGATTTAATATTAC	509
QY	732	TTACAGTTTTCATATGTCTCAATTAATCTCTAGCGCCGAGGAGTGAAGATCAGTAA	791
DB	510	TTACAGTTTTCATATGTCTCAATTAATCTCTAGCGCCGAGGAGTGAAGATCAGTAA	569
QY	792	TAGCAGCGAAACTGTGAATGTGAATGTCTGAAACTGGAAAGGTGAACATGTGACAT	851
DB	570	TAGCAGCGAATCTGTGAAATGTGAATGTCTGAAACTGGAAAGGTGAACATGTGACAT	629
QY	852	TCCTCACTGTACACAACTGTGGTTTTCCTCATCGAGGATCTGCATTTCAAGTGTAT	911
DB	630	TCCTCACTGTACACAACTGTGGTTTTCCTCATCGAGGATCTGCATTTCAAGTGTAT	699
QY	912	CAGAGGATGTCTTCTTCTCAGACTGGCAGGGTCTCTGGATGTTTCACTTCTCTACACG	971
DB	690	CAGAGGATGTCTTCTTCTCAGACTGGCAGGGTCTCTGGATGTTTCACTTCTCTACACG	749
QY	972	TAAACAGTCAATTTTGGACTCGAGAGGAATATTCTAACTTAAAGTCTCCAGAGCATCTCA	1031
DB	750	TAAACAGTCAATTTTGGACTCGAGAGGAATATTCTAACTTAAAGTCTCCAGAGCATCTCA	809
QY	1032	TAAAGCTGTGTCAATGGAACATTAATGTGGTGTGTGGAGGATATATGTTTCAACCACTC	1091
DB	810	TAAAGCTGTGTCAATGGAACATTAATGTGGTGTGTGGAGGATATATGTTTCAACCACTC	869

QY	1092	AGATTATAACATGGTCTTAGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCG	1151
DB	870	AGATTATAACATGGTCTTAGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCG	929
QY	1152	TTCTGTGAACAATGTGGTGTGTAGATATGGTCAATCTTTTGGCATTTATACAGGATAAAT	1211
DB	930	TTCTGTGAACAATGTGGTGTGTAGATATGGTCAATCTTTTGGCATTTATACAGGATAAAT	989
QY	1212	TTACATGTATGGAGAAAAATTTGATCCAACTGGGAATGTGACCAATGATGTGAGAGTTT	1271
DB	990	TTACATGTATGGAGAAAAATTTGATCCAACTGGGAATGTGACCAATGATGTGAGAGTTT	1049
QY	1272	TCACATTCATATGATGATCATGGTGTGTGACCCCTTAAGGCAAAAGGAGCAGTATGCAGT	1331
DB	1050	TCACATTCATATGATGATCATGGTGTGTGACCCCTTAAGGCAAAAGGAGCAGTATGCAGT	1109
QY	1332	GGTTGGGCACTCTGCACACATTTTACACTGNAAGATGGCCGAGTGGTCACTGCTGGTCAT	1391
DB	1110	GGTTGGGCACTCTGCACACATTTTACACTGNAAGATGGCCGAGTGGTCACTGCTGGTCAT	1169
QY	1392	CTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAA	1451
DB	1170	CTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAA	1229
QY	1452	GAACACATGGAGTATATTACACACCCAGGGTCCCTTGTGCAAGGGGTTACGGCCATAG	1511
DB	1230	GAACACATGGAGTATATTACACACCCAGGGTCCCTTGTGCAAGGGGTTACGGCCATAG	1289
QY	1512	CAGTGTTTACGACCACTAGGACCCAGGGCCCTATACGTTTCATGTTGCTCAAGGCTTTT	1571
DB	1290	CAGTGTTTACGACCACTAGGACCCAGGGCCCTATACGTTTCATGTTGCTCAAGGCTTTT	1349
QY	1572	TGCCAATAGTACCCGGTTGCAGATGATCTCTACCGATATGATGTGGATATCCAGATGTG	1631
DB	1350	TGCCAATAGTACCCGGTTGCAGATGATCTCTACCGATATGATGTGGATATCCAGATGTG	1409
QY	1632	GACCATCTTAAAGACACCGCATTTTCCGTTTACTTGCACACAGCTGTGATGTAGTGG	1691
DB	1410	GACCATCTTAAAGACACCGCATTTTCCGTTTACTTGCACACAGCTGTGATGTAGTGG	1469
QY	1692	AACCATGTGTGTGTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCAA	1751
DB	1470	AACCATGTGTGTGTGGGGGAAACACACACAATGACACATCTATGAGCCATGGCGCAA	1529
QY	1752	ATGCTTCTTTCAGATTTTCATGGCCATGACATTTGCCCTGTGACCGCTGGTCACTGCTTC	1811
DB	1530	ATGCTTCTTTCAGATTTTCATGGCCATGACATTTGCCCTGTGACCGCTGGTCACTGCTTC	1589
QY	1812	CAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAGCAGTCTTTACACAACA	1870
DB	1590	CAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTTCAGCAGTCTTTACACAACA	1648
QY	1871	GCACCATGTATGTGTGGTGTTCATATGTCTCTCTCTCCTCCTCAGCGACATCTGATTTCA	1930
DB	1649	GCACCATGTATGTGTGGTGTTCATATGTCTCTCTCTCCTCCTCAGCGACATCTGATTTCA	1708
QY	1931	CCTCGGAAACAGTGTGATCGCATCGAGTGAAGCGCTTGTGTTAGCAGCAGGACTGGTA	1990
DB	1709	CCTCGGAAACAGTGTGATCGCATCGAGTGAAGCGCTTGTGTTAGCAGCAGGACTGGTA	1768
QY	1991	TTCCGTGTGTGGAAACACAGGGTCTCTAGTGTATCTCTGTGGCGCTGGCAACTGATG	2050
DB	1769	TTCCGTGTGTGGAAACACAGGGTCTCTAGTGTATCTCTGTGGCGCTGGCAACTGATG	1828
QY	2051	TACAGAGAAAGCTTTAAATCAGAACTGTTTTTCCAAAGAACTCTTGACCATCACAGAT	2110
DB	1829	TACAGAGAAAGCTTTAAATCAGAACTGTTTTTCCAAAGAACTCTTGACCATCACAGAT	1888
QY	2111	GTGACCAACACACAGATTTGTATACAGCTGTACAGCCAAACACCAATGACCTGGTGCA	2170
DB	1889	GTGACCAACACACAGATTTGTATACAGCTGTACAGCCAAACACCAATGACCTGGTGCA	1948

Qy	2171	ATGACCAATTGTGTCCTCCACAGAACACACAGCTGTCTCAGNAGGCCAGATCTCCATTTTGTAGGT	2233
Db	1949	ATGACCAATTGTGTCCTCCACAGAACACACAGCTGTCTCAGNAGGCCAGATCTCCATTTTGTAGGT	2008
Qy	2231	ATGAGAATTGCCCCACAGGATAACCCCATGTACTACTGTAAACAAGAAAGACAGCTGCAGGA	2290
Db	2009	ATGAGAATTGCCCCACAGGATAACCCCATGTACTACTGTAAACAAGAAAGACAGCTGCAGGA	2068
Qy	2291	GCTGTGCCCTGGACACAGAACTGCGCAGTGGAGGCCCGGAAATCAGAGAGTGCAATTGCCCTGCG	2350
Db	2069	GCTGTGCCCTGGACACAGAACTGCGCAGTGGAGGCCCGGAAATCAGAGAGTGCAATTGCCCTGCG	2128
Qy	2351	CCGAAATAATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAATAATTACTTA	2410
Db	2129	CCGAAATAATCTGTGGCATTGGCTGGCATTTGGTTGGAAACTCATGTTTGAAATAATTACTTA	2188
Qy	2411	CTGCCAAGGAGAAATTAGCAATGCTCTAAATTTGTTCTGTAGGAACCAAAATGCCCTTTTGG	2470
Db	2189	CTGCCAAGGAGAAATTAGCAATGCTCTAAATTTGTTCTGTAGGAACCAAAATGCCCTTTTGG	2248
Qy	2471	CTTCTCTTACAAACCCAGAGAGGTAGAAATTTCTTCCTTAAGCAGCTGCGGAATAATGCAGT	2530
Db	2249	CTTCTCTTACAAACCCAGAGAGGTAGAAATTTGTTCTTAAGCAGCTGCGGAATAATGCAGT	2308
Qy	2531	CATCTCAGAGCATGTCCAAAGCTCACCTTAAACCCCATGGGTGCGCCTTCGGAAGATCAATG	2590
Db	2309	CATCTCAGAGCATGTCCAAAGCTCACCTTAAACCCCATGGGTGCGCCTTCGGAAGATCAATG	2368
Qy	2591	TGTCCTACTGGTGGGAGAGATATGTCCCATTTACAAATAGTTTACTACAGTGATGC	2650
Db	2369	TGTCCTACTGGTGGGAGAGATATGTCCCATTTACAAATAGTTTACTACAGTGATGC	2428
Qy	2651	CGTCTGAGCCCAAGTGATGCTGGAAATTCGTGGAAATTTTATCAGAAACCAAGTACTCGGGGAC	2710
Db	2429	CGTCTGAGCCCAAGTGATGCTGGAAATTCGTGGAAATTTTATCAGAAACCAAGTACTCGGGGAC	2488
Qy	2711	TGAAGGCTGCAACTGTGCAATCAACCCACTCAATGGTAGTGTCGTGAAAGGCCCTGCAAAACC	2770
Db	2489	TGAAGGCTGCAACTGTGCAATCAACCCACTCAATGGTAGTGTCGTGAAAGGCCCTGCAAAACC	2548
Qy	2771	ACAGTGCTAAGCAGTGC CGGACACCACTGTCCTTGAGGACAGCATGTGGAGATTGCACCA	2830
Db	2549	ACAGTGCTAAGCAGTGC CGGACACCACTGTCCTTGAGGACAGCATGTGGAGATTGCACCA	2608
Qy	2831	CGCGCAGCTCTGAGTGCAATGTGTGTCAGCAACATGAAGCAGTGTGTGGACTCCAATGCGCT	2890
Db	2609	CGCGCAGCTCTGAGTGCAATGTGTGTCAGCAACATGAAGCAGTGTGTGGACTCCAATGCGCT	2668
Qy	2891	ATGTGGCCTCTTTCCTTTTGGCCAGTGATGGAATGGTATACGATGAGCACCTGTGCCCTC	2950
Db	2669	ATGTGGCCTCTTTCCTTTTGGCCAGTGATGGAATGGTATACGATGAGCACCTGTGCCCTC	2728
Qy	2951	CTGAAAATTTGTACAGGCTACTGTACTCTGTAGTCAATTCCTTTGGAGCAACCAAGGCTGTGGCT	3010
Db	2729	CTGAAAATTTGTACAGGCTACTGTACTCTGTAGTCAATTCCTTTGGAGCAACCAAGGCTGTGGCT	2788
Qy	3011	GGTGTACTGATCCACGCAATCTGGCAAAAGGAAATGCATAGAGGGTTCTCTATAAAGGAC	3070
Db	2789	GGTGTACTGATCCACGCAATCTGGCAAAAGGAAATGCATAGAGGGTTCTCTATAAAGGAC	2848
Qy	3071	CAGTGAAGATGCCCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTCTGCTCAATT	3130
Db	2849	CAGTGAAGATGCCCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTCTGCTCAATT	2908
Qy	3131	CCAGCATGTGTCTAGAGGACGACAGATACAACTGGTCTTTTCAATTCAGTGTCCAGCTTGCC	3190
Db	2909	CCAGCATGTGTCTAGAGGACGACAGATACAACTGGTCTTTTCAATTCAGTGTCCAGCTTGCC	2968
Qy	3191	AATCCACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAAAGTGTGAGAACTCTGA	3250
Db	2969	AATCCACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAAAGTGTGAGAACTCTGA	3028
Qy	3251	CCACAGCAAGCATCGCGAGACCTGCATATCTGTGCTTCTACGGTGTCTACGGTGTATCCCAACCAATGGAG	3310

## RESULT 8

AAS72657

AA372037  
ID AAS72657 standard: cDNA: 3597 BP.

XX  
AC  
78573657.

13-FEB-2003 (five+ ent v)

[illegible]

XX

kw food supplement; medical imaging; diagnostic; genetic disorder;

XX

XX

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

XX  
XX  
1002-100-11

30-MAR-2001; 2001WC-03008631.  
PF  
XX

PR 31-MAR-2000; 2000US-00540217.  
PB 31-MAR-2000; 2000US-00548157

U  
Z  
H  
  
C  
P  
O  
T  
E  
S  
  
-  
D  
C  
T  
E  
S  
  
X

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

[illegible]

DR P-PSDB; ABG08470.

PT New isolated poly

PI DIAGNOSTICS, TOEMLERES, GERE MAPPIING, THERMOCALCULATION OF "MACROCALCULATION"

PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

PS Claim 1; SEQ ID NO 8461; 103bp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activities. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 71.0%; Score 3045; DB 5; Length 3597;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3485; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

QY	312	CTGTGTCAACCGCGTGCCTGCACACCTGCGACCGGCGAGTGGCTGCCCCCGCGCTG	371
DB	90	CTGTGTCAACCGCGTGCCTGCACACCTGCGACCGGCGAGTGGCTGCCCCCGCGCTG	149
QY	372	GGTGGGCGAGCAATGCCAGCACTCGGGGGCGGCTTCAGACTAACTGGATCTCTGGGTT	431
DB	150	GGTGGGCGAGCAATGCCAGCACTCGGGGGCGGCTTCAGACTAACTGGATCTCTGGGTT	209
QY	432	TGTGACAGATGGACTGGAAATTAATAACAAACGAGTGCAGTGGCTCAATGAAGG	491
DB	210	TGTGACAGATGGACTGGAAATTAATAACAAACGAGTGCAGTGGCTCAATGAAGG	269
QY	492	ACAGCCAAATAGAAATATGAGACTTCGTTCAATCATTTTGCTACAGAGTGTAGTTGGGA	551
DB	270	ACAGCCAAATAGAAATATGAGACTTCGTTCAATCATTTTGCTACAGAGTGTAGTTGGGA	329
QY	552	CCATTTATGTTTATGATGGGACTCAATTTATGCAACCGCTAGTGTCTGCATTTAGTGG	611
DB	330	CCATTTATGTTTATGATGGGACTCAATTTATGCAACCGCTAGTGTCTGCATTTAGTGG	389
QY	612	CCTCATTTCTCTGAGAGATGCAATGAGACTGCTCCCTGAGGTTGTCGCACATCAGG	671
DB	390	CCTCATTTCTCTGAGAGATGCAATGAGACTGCTCCCTGAGGTTGTCGCACATCAGG	449
QY	672	TTATGCGCTTGTGCAATTTTTTATGATGCTGCTTATAATTTGACTGGATTTAATATTAC	731
DB	450	TTATGCGCTTGTGCAATTTTTTATGATGCTGCTTATAATTTGACTGGATTTAATATTAC	509
QY	732	TTACAGTTTGTATGTGTCAAATTAATCTGCTAGGCGGAGAGAGTGAAGTACAGTAA	791
DB	510	TTACAGTTTGTATGTGTCAAATTAATCTGCTAGGCGGAGAGAGTGAAGTACAGTAA	569
QY	792	TAGCAGCGAAACTGTTGAAATGTGTAATCTCTGAAACTGGAAGTGAACATGTGACAT	851
DB	570	TAGCAGCGAATCTGTTGAAATGTGTAATCTCTGAAACTGGAAGTGAACATGTGACAT	629
QY	852	TCCTCACTGTACAGACAACTGTGTTTTTCTCATCGAGGGCATCTGCAATTCAGATGT	911
DB	630	TCCTCACTGTACAGACAACTGTGTTTTTCTCATCGAGGGCATCTGCAATTCAGATGT	689

QY	912	CAAGAGATGCTCTCTGCTTCTCAGACTGGCAGGGTCTCGATGTTCTCAGTTCCTGTACCAAGC	971
DB	690	CAGAGGATGCTCTCTGCTTCTCAGACTGGCAGGGTCTCGATGTTCTCAGTTCCTGTACCAAGC	749
QY	972	TAAACAGTCAATTTTGGACTCGAGAGGAATATTTCTAACTTAAAGCTCCCGACAGCATCTCA	1031
DB	750	TAAACAGTCAATTTTGGACTCGAGAGGAATATTTCTAACTTAAAGCTCCCGACAGCATCTCA	809
QY	1032	TAAAGCTGTGTCATATGGAACCAATTTATGTGGTCTGTTGGAGATATATGTTCAACCACTC	1091
DB	810	TAAAGCTGTGTCATATGGAACCAATTTATGTGGTCTGTTGGAGATATATGTTCAACCACTC	869
QY	1092	AGATATATACATGTTCTTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTTAAACCG	1151
DB	870	AGATATATACATGTTCTTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTTAAACCG	929
QY	1152	TTCTGTGAACCAATGTGGTGTGTAGATATGCTCATTTCTTTGGCATTTATCAAGGATAAAT	1211
DB	930	TTCTGTGAACCAATGTGGTGTGTAGATATGCTCATTTCTTTGGCATTTATCAAGGATAAAT	989
QY	1212	TTACATGTATGGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGATGATGAGATTTT	1271
DB	990	TTACATGTATGGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGATGATGAGATTTT	1049
QY	1272	TCACATTCATATGATGCTGCTGTTGACCCCTAAGCCAAAGGAGGAGTATGCAGT	1331
DB	1050	TCACATTCATATGATGCTGCTGTTGACCCCTAAGCCAAAGGAGGAGTATGCAGT	1109
QY	1332	GGTTGGGCACTCTGCACACATTTTACACTGGAAGATGGCCGAGTGTCTGCTGCTCAT	1391
DB	1110	GGTTGGGCACTCTGCACACATTTTACACTGGAAGATGGCCGAGTGTCTGCTGCTCAT	1169
QY	1392	CTTTGCTCATCTGCCCTCTCTATGGAATATTAAGCAATGTGAGGAATATGATTTGGATAA	1451
DB	1170	CTTTGCTCATCTGCCCTCTCTATGGAATATTAAGCAATGTGAGGAATATGATTTGGATAA	1229
QY	1452	GAACACATGGAGTATATTACACACCCAGGGTCCCTTGTGCAAGGGGTACGCGCCATAG	1511
DB	1230	GAACACATGGAGTATATTACACACCCAGGGTCCCTTGTGCAAGGGGTACGCGCCATAG	1289
QY	1512	CAGTGTTCAGACCATAGACCCGCGCCCTATAGCTTCATGCTGGCTACAGGCTTTTACG	1571
DB	1290	CAGTGTTCAGACCATAGACCCGCGCCCTATAGCTTCATGCTGGCTACAGGCTTTTACG	1349
QY	1572	TGCCAATTAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGGAATACCCAGATGTG	1631
DB	1350	TGCCAATTAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGGAATACCCAGATGTG	1409
QY	1632	GACCATTTCTTAAGGACAGCCGATTTTCCGTTACTTGCACACAGCTGTGATGAGTGG	1691
DB	1410	GACCATTTCTTAAGGACAGCCGATTTTCCGTTACTTGCACACAGCTGTGATGAGTGG	1469
QY	1692	AACCATGCTGCTGTTTGGGGGAAACACACACATGACACATCTATGACCCATGCGCCAA	1751
DB	1470	AACCATGCTGCTGTTTGGGGGAAACACACACATGACACATCTATGACCCATGCGCCAA	1529
QY	1752	ATGCTTCTCTCAGATTTTATGGGCTATGACATTTGCCCTGTGACCGCTGGTCACTGCTTC	1811
DB	1530	ATGCTTCTCTCAGATTTTATGGGCTATGACATTTGCCCTGTGACCGCTGGTCACTGCTTC	1589
QY	1812	CAGACCTCATCTCCACCATGATGTCAACAGATTTGGCCATTT-CAGCAGTCTTTACACACA	1870
DB	1590	CAGACCTCAT-TCCACCATGATGTCAACAGATTTGGCCATTTCCAGCAGTCTTTACACACA	1648
QY	1871	GCACCATGATGTGTTCGGTGTGTTTCAATAGTCTCTCTCCTCAGCEACATCTGATATCA	1930
DB	1649	GCACCATGATGTGTTCGGTGTGTTTCAATAGTCTCTCTCCTCAGCEACATCTGATATCA	1708
QY	1931	CCTCGGAACAGTGTGATCGGCATCGGAGTGAAGCGCTTGTGTTAGCAGCAGCCTGGTA	1990
DB	1709	CCTCGGAACAGTGTGATCGGCATCGGAGTGAAGCGCTTGTGTTAGCAGCAGCCTGGTA	1768
QY	1991	TTCCGTTGTGTGGAAACACACAGGGTCTGCTCAGTGTATCTCGTGGGCGCTGCCAATGATG	2050

Db 1769 TTGGTGTGTGGAAACACAGGGTCGTCTCAGTGTATCTCTGTGGCGCTGGCACTGATG 1828  
QY 2051 AACAGAAAGAAAGTTTAAATCAGAAATGTTTTCAAAAGAACTCTTTGACCATGACAGAT 2110  
Db 1829 AACAGAAAGAAAGTTTAAATCAGAAATGTTTTCAAAAGAACTCTTTGACCATGACAGAT 1888  
QY 2111 GTGACACGACACACAGATTTGTTTACAGCTGTGACAGCAACACCAATGACTGCCACCTGGTGCA 2170  
Db 1889 GTGACACGACACACAGATTTGTTTACAGCTGTGACAGCAACACCAATGACTGCCACCTGGTGCA 1948  
QY 2171 ATGACCAATTTGTTCCCGAGGAACACACAGCTGCTCAGAAAGGCAGATCTCCATTTTATAGT 2230  
Db 1949 ATGACCAATTTGTTCCCGAGGAACACACAGCTGCTCAGAAAGGCAGATCTCCATTTTATAGT 2008  
QY 2231 ATGAGAAATTTGCCCAAGATAACCCATGACTACTGTATCAAGAACACAGCTGACAGGA 2290  
Db 2009 ATGAGAAATTTGCCCAAGATAACCCATGACTACTGTATCAAGAACACAGCTGACAGGA 2068  
QY 2291 GCTGTGCCCTGGACACAGAACTGCCAGTGGGAGGCCCGGAATCAGAGTGCATTGCCCTGC 2350  
Db 2069 GCTGTGCCCTGGACACAGAACTGCCAGTGGGAGGCCCGGAATCAGAGTGCATTGCCCTGC 2128  
QY 2351 CCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACCTCATGTTGAAATTTACTA 2410  
Db 2129 CCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACCTCATGTTGAAATTTACTA 2188  
QY 2411 CTGCCAAGGAAATATCACAATGCTAAATGTTCTGTAGGAACCAACAATGCCCTTTGG 2470  
Db 2189 CTGCCAAGGAAATATCACAATGCTAAATGTTCTGTAGGAACCAACAATGCCCTTTGG 2248  
QY 2471 CTTCTCTTAAACCCAGAAAGTAGAATTTGCTTTAGCAGCTGCGGAATATGACGT 2530  
Db 2249 CTTCTCTTAAACCCAGAAAGTAGAATTTGCTTTAGCAGCTGCGGAATATGACGT 2308  
QY 2531 CATCTCAGAGCATGTCGAAGCTCACTTAAACCCATGGTGGCTTTCGGAAGATCAATG 2590  
Db 2309 CATCTCAGAGCATGTCGAAGCTCACTTAAACCCATGGTGGCTTTCGGAAGATCAATG 2368  
QY 2591 TGTCTACTGTGTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGC 2650  
Db 2369 TGTCTACTGTGTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGC 2428  
QY 2651 CGTCTGAGCCAGTATGCTGATCTGTGGAATTTTATCAGAACCCAGTACTCGGGAC 2710  
Db 2429 CGTCTGAGCCAGTATGCTGATCTGTGGAATTTTATCAGAACCCAGTACTCGGGAC 2488  
QY 2711 TGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCTGCAAAAC 2770  
Db 2489 TGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCTGCAAAAC 2548  
QY 2771 ACAGTGTAAAGCAGTGGCGGACACCATGTGCTTGGAGACAGCATGTGGAGATTGCACCA 2830  
Db 2549 ACAGTGTAAAGCAGTGGCGGACACCATGTGCTTGGAGACAGCATGTGGAGATTGCACCA 2608  
QY 2831 GCGGCAGCTCTGAGTGCATGTGGTGACAGCAACATGAGCAGTGTGGACTCCCAATGCCT 2890  
Db 2609 GCGGCAGCTCTGAGTGCATGTGGTGACAGCAACATGAGCAGTGTGGACTCCCAATGCCT 2668  
QY 2891 ATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGTATACGATGAGCACTGCCCCC 2950  
Db 2669 ATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGTATACGATGAGCACTGCCCCC 2728  
QY 2951 CTGAAATTTGTTCCAGCTACTGTACTCTGATCTTGGAGCAACCGAGCTGTGGCT 3010  
Db 2729 CTGAAATTTGTTCCAGCTACTGTACTCTGATCTTGGAGCAACCGAGCTGTGGCT 2788  
QY 3011 GGTGTACTGATCCAGCAATATCGCAAGGAAATGATAGAGGGTTCTCTATTAAGGAC 3070  
Db 2789 GGTGTACTGATCCAGCAATATCGCAAGGAAATGATAGAGGGTTCTCTATTAAGGAC 2848  
QY 3071 CAGTGAAGATGCTTCGCAAGCCCTCAGGAAATTTCTATCCACAGCCCTGCTCAATT 3130

Db 2849 CAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATT 2908  
QY 3131 CAGCATGTGTCTAGAGGACACAGATCAAACTGGTCTTTTCAATTCATCTGTCAGGTTGCC 3190  
Db 2909 CAGCATGTGTCTAGAGGACACAGATCAAACTGGTCTTTTCAATTCATCTGTCAGGTTGCC 2968  
QY 3191 AATGCAACGGCCACAGTAATATCATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGA 3250  
Db 2969 AATGCAACGGCCACAGTAATATCATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGA 3028  
QY 3251 CCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCAATGGAG 3310  
Db 3029 CCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCAATGGAG 3088  
QY 3311 GGAATATGTCCAGCATGCAAGTGAAGTGGCAAGCGCTCTCTGTGCAACCAACACAGGGCA 3370  
Db 3089 GGAATATGTCCAGCATGCAAGTGAAGTGGCAAGCGCTCTCTGTGCAACCAACACAGGGCA 3148  
QY 3371 AGTGCTTCTGCAACCAAGGGGCTCAAGGGGGAAGAGTCCAGCTATGTGAGGTAGAAA 3430  
Db 3149 AGTGCTTCTGCAACCAAGGGGCTCAAGGGGGAAGAGTCCAGCTATGTGAGGTAGAAA 3208  
QY 3431 ATCGATACCAAGGAAACCTCTCAGAGGAAATGTTATTTATATCTCTTCTTATTGACTATC 3490  
Db 3209 ATCGATACCAAGGAAACCTCTCAGAGGAAATGTTATTTATATCTCTTCTTATTGACTATC 3268  
QY 3491 AGTTCACCTTTAGTCTATCCAGGAAGATGCTGCTATTTATACACAGCTATCAATTTTGTGG 3550  
Db 3269 AGTTCACCTTTAGTCTATCCAGGAAGATGCTGCTATTTATACACAGCTATCAATTTTGTGG 3328  
QY 3551 CTACTCTCAGCAACCAACAGGATTTGGACATGTTTCATCAATGCTCCAAAGATTTCA 3610  
Db 3329 CTACTCTCAGCAACCAACAGGATTTGGACATGTTTCATCAATGCTCCAAAGATTTCA 3388  
QY 3611 ACCTCAACATCACTGGGCTGCGAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGC 3670  
Db 3389 ACCTCAACATCACTGGGCTGCGAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGC 3448  
QY 3671 CTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTTG 3730  
Db 3449 CTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTTG 3508  
QY 3731 ATTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTTCACTGGCCCA 3790  
Db 3509 ATTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTTCACTGGCCCA 3568  
QY 3791 TCAAAATTCAG 3801  
Db 3569 TCAAAATTCAG 3579

RESULT: 9  
ADD70997  
ID ADD70997 standard; DNA; 3597 BP.  
XX  
AC ADD70997;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
Human attractin gene SEQ ID NO:1.  
DE  
XX  
XX liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;  
XX cytotstatic; gene therapy; human; gene; ds.  
XX  
OS Homo sapiens.  
XX  
FN WO2003061564-A2.  
PD 31-JUL-2003.  
XX  
XX 20-DEC-2002; 2002WO-US040718.  
PF  
XX 21-DEC-2001; 2001US-0341815P.  
PR

PR 31-DEC-2001; 2001US-0343185P.  
XX (GENE-) GENE LOGIC INC.  
PA (LGI-) LG BIOMEDICAL INST.  
XX  
XX Koh SS, Liu Q, Chung H, Zeng W, Lee B, Yeramilli S, Song SY;  
XX WPI; 2003-663343/62.  
XX  
XX Diagnosing liver cancer cells, useful for treating liver cancer  
PT associated with chronic hepatitis or cirrhosis comprises detecting the  
PT level of expression in a tissue sample of one or more genes associated  
PT with cancerous liver tissues.  
XX  
XX Claim 1; SEQ ID NO 1; 176pp; English.  
XX  
XX The present invention describes a method for diagnosing liver cancer  
CC cells comprising detecting the level of expression in a tissue sample of  
CC one or more genes given in the specification (see ADD70997 to ADD71105),  
CC where differential expression of the genes is indicative of liver cancer.  
CC Also described: (1) detecting the progression of liver cancer in a  
CC patient; (2) monitoring the treatment of a patient with liver cancer; (3)  
CC treating a patient with liver cancer; (4) typing a liver disease in a  
CC patient; (5) detecting the presence or progression of liver cancer in a  
CC patient with chronic hepatitis or cirrhosis; (6) differentiating liver  
CC cancer related to chronic hepatitis from liver cancer related to  
CC cirrhosis; (7) screening for an agent capable of modulating the onset or  
CC progression of liver cancer; (8) a composition comprising at least two  
CC oligonucleotides comprising a sequence that specifically hybridizes to  
CC any of the genes; (9) a solid support comprising the at least two  
CC oligonucleotides; (10) a computer system comprising a database containing  
CC information identifying the level in liver tissue of a set of genes; (11)  
CC a method for using the computer system to present information identifying  
CC the expression level in tissue or cell of any of the genes; and (12) a  
CC therapeutic agent for slowing or halting the progression of liver cancer.  
CC The methods are useful for treating liver cancer associated with chronic  
CC hepatitis or cirrhosis. The present sequence represents a specifically  
CC claimed human gene sequence which is used in the exemplification of the  
CC present invention.  
XX  
XX Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;  
XX  
XX Query Match 71.0%; Score 3045; DB 9; Length 3597;  
XX Best Local Similarity 99.8%; Pred. No. 0;  
XX Matches 3485; Conservative 0; Mismatches 4; Indels 2; Gaps 2;  
XX  
QY 312 CTGTGTCACCGCGTCTGCAACCCCTGGACCGCGGCGAGTGCCTGCCCCCGCGCTG 371  
DB 90 CTGTGTCACCGCGTCTGCAACCCCTGGACCGCGGCGAGTGCCTGCCCCCGCGCTG 149  
QY 372 GGTGGGCGAGCAATCCAGCACTCGGGGGCGGCTTCAGACTACTGGATCTCTCGGTT 431  
DB 150 GGTGGGCGAGCAATCCAGCACTCGGGGGCGGCTTCAGACTACTGGATCTCTCGGTT 209  
QY 432 TGTGACAGATGACCTGGAAATTTAAATACAAACGAAAGTGCACGTGGCTCATTTGAAGG 491  
DB 210 TGTGACAGATGACCTGGAAATTTAAATACAAACGAAAGTGCACGTGGCTCATTTGAAGG 269  
QY 492 ACAGCAATAGAAATAGCACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTTGGGA 551  
DB 270 ACAGCAATAGAAATAGCACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTTGGGA 329  
QY 552 CCAATTATATGTTATGATGGGACTCAATTTATGACCGGCTAGTGTGCTGCAATTAGTGG 611  
DB 330 CCAATTATATGTTATGATGGGACTCAATTTATGACCGGCTAGTGTGCTGCAATTAGTGG 389  
QY 612 CCTCATTTGCTGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGG 671  
DB 390 CCTCATTTGCTGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGG 449  
QY 672 TTATGCCCTTGTGCAATTTTTTTAGTGATGCGCTTATTAATTTGACTGGATTTAATATTAC 731  
DB 450 TTATGCCCTTGTGCAATTTTTTTAGTGATGCGCTTATTAATTTGACTGGATTTAATATTAC 509

QY 732 TTACAGTTTTTGATATGTGTCCAAATAACTGCTCAGSCCGAGGAGAGTGTAAAGATCAGTAA 791  
DB 510 TTACAGTTTTTGATATGTGTCCAAATAACTGCTCAGSCCGAGGAGAGTGTAAAGATCAGTAA 569  
QY 792 TAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAATCTGAAAGAGTGAAGCATGTGACAT 851  
DB 570 TAGCAGCGAATCTGTTGAATGTGAATGTTCTGAAATCTGAAAGAGTGAAGCATGTGACAT 629  
QY 852 TCCTCACTGTACAGACAACCTGTTTCTCTCATCGAGGCATCTGCAATTCAGTGTATGT 911  
DB 630 TCCTCACTGTACAGACAACCTGTTTCTCTCATCGAGGCATCTGCAATTCAGTGTATGT 689  
QY 912 CAGAGGATGCTCTGCTTCTCTCAGACTGCGCAGGCTCTCGATGTTCTGATCTCTGATCCAGC 971  
DB 690 CAGAGGATGCTCTGCTTCTCTCAGACTGCGCAGGCTCTCGATGTTCTGATCTCTGATCCAGC 749  
QY 972 TAACAGTCAATTTGAGCTGCGAGAGAAATATCTTAACCTTAAAGCTCCCGAGAGCATCTCA 1031  
DB 750 TAACAGTCAATTTGAGCTGCGAGAGAAATATCTTAACCTTAAAGCTCCCGAGAGCATCTCA 809  
QY 1032 TAAAGCTGTGGTCAATGGAACCAATATGTTGGTGTGAGGATATATGTTCAACCACTC 1091  
DB 810 TAAAGCTGTGGTCAATGGAACCAATATGTTGGTGTGAGGATATATGTTCAACCACTC 869  
QY 1092 AGATTATPAACTGGTCTTAGCGTATGACCTTCTCTAGGGAGTGGCTTCCACTAAACCG 1151  
DB 870 AGATTATPAACTGGTCTTAGCGTATGACCTTCTCTAGGGAGTGGCTTCCACTAAACCG 929  
QY 1152 TTCTGTGCAATGTTGTTTAGATATGTTGATCTATCTTTTGGCATTTACAAGGATAAAT 1211  
DB 930 TTCTGTGCAATGTTGTTTAGATATGTTGATCTATCTTTTGGCATTTACAAGGATAAAT 989  
QY 1212 TTACATGTATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGATGTGAGATTTT 1271  
DB 990 TTACATGTATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGATGTGAGATTTT 1049  
QY 1272 TCACATTTCAATGATGATGCTGTTGTTGATGCTTGTGATGCTTGTGATGCTTGTGATGCT 1331  
DB 1050 TCACATTTCAATGATGATGCTGTTGTTGATGCTTGTGATGCTTGTGATGCTTGTGATGCT 1109  
QY 1332 GGTGGGCACTCTGCACACATTTGTACACTGAAGAAATGCCGAGTGGTCTGCTGCTCAT 1391  
DB 1110 GGTGGGCACTCTGCACACATTTGTACACTGAAGAAATGCCGAGTGGTCTGCTGCTCAT 1169  
QY 1392 CTTTGGTCTACTGCGCTCTCTATGAGATATATAGCAATGTGAGGAAATATGATTTGGATAA 1451  
DB 1170 CTTTGGTCTACTGCGCTCTCTATGAGATATATAGCAATGTGAGGAAATATGATTTGGATAA 1229  
QY 1452 GAACACATGAGTATATACACACCCAGGCTGCTTGTGCAAGGGGTGTACGGCCATAG 1511  
DB 1230 GAACACATGAGTATATACACACCCAGGCTGCTTGTGCAAGGGGTGTACGGCCATAG 1289  
QY 1512 CAGTGTGTTACGACCATAGGACCCAGGCGCTTACGTTTATGTTGCTCAAGGCTTTTCCAG 1571  
DB 1290 CAGTGTGTTACGACCATAGGACCCAGGCGCTTACGTTTATGTTGCTCAAGGCTTTTCCAG 1349  
QY 1572 TGCCATAGTACCGCTTGCAGATGATCTCTACCGATATGATGTTGATGATGATGATGATGATG 1631  
DB 1350 TGCCATAGTACCGCTTGCAGATGATCTCTACCGATATGATGTTGATGATGATGATGATGATG 1409  
QY 1632 GACCATTTCTTAAGGACAGCCGATTTTTTCCGTTTACTTGCACACAGCTGTGATGTTGAGTGG 1691  
DB 1410 GACCATTTCTTAAGGACAGCCGATTTTTTCCGTTTACTTGCACACAGCTGTGATGTTGAGTGG 1469  
QY 1692 AACCATGCTGGTGTGTTGGGGGAAACACACAAATGACACATCTATGAGCCATGGGCCAA 1751  
DB 1470 AACCATGCTGGTGTGTTGGGGGAAACACACAAATGACACATCTATGAGCCATGGGCCAA 1529  
QY 1752 ATGCTTCTCTCAGATTTTCATGGCCCTATGACATGCTGTCACCGCTGTCAGTGTGCTTCC 1811  
DB 1530 ATGCTTCTCTCAGATTTTCATGGCCCTATGACATGCTGTCACCGCTGTCAGTGTGCTTCC 1589

[illegible]



XX Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;  
 KW weight regulation; cell therapy; body weight disorder; cachexia;  
 KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
 KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.  
 XX Homo sapiens.  
 XX WO200005373-A2.  
 XX 03-FEB-2000.  
 XX 21-JUL-1999; 99WO-US016484.  
 XX 21-JUL-1998; 98US-0093630P.  
 XX 20-OCT-1998; 98US-0104978P.  
 XX 05-FEB-1999; 99US-00245041.  
 XX (MILL-) MILLENIUM PHARM INC.  
 XX Moore K, Nagle DL;  
 XX WPI; 2000-195103/17.  
 XX P-PSDB; AAY81809.  
 XX New human and murine mahogany genes, useful, e.g. for diagnosis and  
 PT treatment of body weight disorders.  
 XX Claim 1; Fig 20a; 18pp; English.  
 XX This sequence represents a human mahogany gene of the invention. The  
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iii) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in (cell-  
 CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (II) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (II) (for diagnosis, prognosis or monitoring of treatment); to  
 CC evaluate (II)-expressing cells intended for cell therapy, and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity. (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other  
 CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate  
 XX  
 XX Sequence 2625 BP; 660 A; 605 C; 694 G; 666 T; 0 U; 0 Other;  
 SQ

Query Match 42.8%; Score 1837; DB 3; Length 2625;  
 Best Local Similarity 99.8%; Pred. No. 0;  
 Matches 2037; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 313 TGTGTCAACGGCGTGGTGGACCTGGACCGCGGCGGCTGCGCGCGCGGCTGG 372  
 DB 310 TGTGTCAACGGCGTGGTGGACCTGGACCGCGGCGGCTGCGCGCGGCTGG 369  
 QY 373 GTGGCGGAGCAATGCCAGCACTGCGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTTT 432  
 DB 370 GTGGCGGAGCAATGCCAGCACTGCGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTTT 429  
 QY 433 GTGACAGATGGACCTGGAAATTTAATAACAAAGAGTGCAGTGGCTCAATTGAAGGA 492  
 DB 430 GTGACAGATGGACCTGGAAATTTAATAACAAAGAGTGCAGTGGCTCAATTGAAGGA 489  
 QY 493 CAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTTGGAC 552  
 DB 490 CAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTTGGAC 549  
 QY 553 CATTATATGTTTATGATGGGAGCTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGC 612

DB 550 CATTATATGTTTATGATGGGAGCTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGC 609  
 QY 613 CTCATTTCTTCTGAGAGAGATGGCAATGACACTGTCCCTGAGGTTGTTGCCACATCAGGT 672  
 DB 610 CTCATTTCTTCTGAGAGAGATGGCAATGACACTGTCCCTGAGGTTGTTGCCACATCAGGT 669  
 QY 673 TATGCCCTTCTGCTGCAATTTTCTAGTGATGCTGCTTATATTTGACTTGACTTAATATCT 732  
 DB 670 TATGCCCTTCTGCTGCAATTTTCTAGTGATGCTGCTTATATTTGACTTGACTTAATATCT 729  
 QY 733 TACAGTTTTCATATGTTCTCCAAATAACTGCTCAGGCGGAGGAGAGTGAAGATCAGTAAT 792  
 DB 730 TACAGTTTTCATATGTTCTCCAAATAACTGCTCAGGCGGAGGAGAGTGAAGATCAGTAAT 789  
 QY 793 AGCAGCGAAACTGTTGAATGTGAATGTTCTGAAAACTGGAAGGTGAAGATGTGACATT 852  
 DB 790 AGCAGCGAATCTGTTGAATGTGAATGTTCTGAAAACTGGAAGGTGAAGATGTGACATT 849  
 QY 853 CCTCAGCTGTACAGCAAACTGTGTTTTCTCTCATCGAGGCACTCTGCAATTCAAGTGATGTC 912  
 DB 850 CCTCAGCTGTACAGCAAACTGTGTTTTCTCTCATCGAGGCACTCTGCAATTCAAGTGATGTC 909  
 QY 913 AGAGGATGCTCCTGCTTCTCAGACTGGCAGGCTCTGGATGTTTCACTTCTGTACAGCT 972  
 DB 910 AGAGGATGCTCCTGCTTCTCAGACTGGCAGGCTCTGGATGTTTCACTTCTGTACAGCT 969  
 QY 973 AACCACTGTTTGGACTCGAGAGGAATATCTTAACCTTAAGCTCCCGAGAGCACTCAT 1032  
 DB 970 AACCACTGTTTGGACTCGAGAGGAATATCTTAACCTTAAGCTCCCGAGAGCACTCAT 1029  
 QY 1033 AAAGCTGTGCTCAATCGAAACCAATTTATGTTGGTGTGTTGGAGGATATATGTTCAACCACTCA 1092  
 DB 1030 AAAGCTGTGCTCAATCGAAACCAATTTATGTTGGTGTGTTGGAGGATATATGTTCAACCACTCA 1089  
 QY 1093 GATTATAACATGTTCTAGGATGACCTTCTTAGGGAGTGGCTTCCACATAACCGT 1152  
 DB 1090 GATTATAACATGTTCTAGGATGACCTTCTTAGGGAGTGGCTTCCACATAACCGT 1149  
 QY 1153 TCTGTGAAACAATGTGGTTGTAGATATGTCATCTTTGGCATTATACAAAGATATAAAT 1212  
 DB 1150 TCTGTGAAACAATGTGGTTGTAGATATGTCATCTTTGGCATTATACAAAGATATAAAT 1209  
 QY 1213 TACATGTATGGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGAGTTGAGATTTT 1272  
 DB 1210 TACATGTATGGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGAGTTGAGATTTT 1269  
 QY 1273 CACATTCATAATGAGTCATGGGTGTTGTTGACCCCTTAAGGCAAAAGAGCAGTATCAGTG 1332  
 DB 1270 CACATTCATAATGAGTCATGGGTGTTGTTGACCCCTTAAGGCAAAAGAGCAGTATCAGTG 1329  
 QY 1333 GTTGGCACTCTGCAACACATTTTACACTGAAGAATGGCCGAGTGGTCAATGCTGTGATC 1392  
 DB 1330 GTTGGCACTCTGCAACACATTTTACACTGAAGAATGGCCGAGTGGTCAATGCTGTGATC 1389  
 QY 1393 TTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAG 1452  
 DB 1390 TTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAG 1449  
 QY 1453 AACACATGGAGTATATTAACACACCCAGGGTGGCCCTTTGTCAGAGGGGTTTACGCCATAGC 1512  
 DB 1450 AACACATGGAGTATATTAACACACCCAGGGTGGCCCTTTGTCAGAGGGGTTTACGCCATAGC 1509  
 QY 1513 AGTGTTTACGACGATAGGACGAGGCCCTTACAGTTTCACTGGTGGCTACAGGCTTTTCACT 1572  
 DB 1510 AGTGTTTACGACGATAGGACGAGGCCCTTACAGTTTCACTGGTGGCTACAGGCTTTTCACT 1569  
 QY 1573 GCCAATAAGTACCGGCTTGCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGG 1632  
 DB 1570 GCCAATAAGTACCGGCTTGCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGG 1629  
 QY 1633 ACCATTCTTAAGGACAGCGGATTTTCCGTTACTTCTTGCAACAGCTGTGATAGTGTGGA 1692  
 DB 1630 ACCATTCTTAAGGACAGCGGATTTTCCGTTACTTCTTGCAACAGCTGTGATAGTGTGGA 1689





Db 983 CATCAATCAGACATCTGTGAGAGTGTGAGACCTGACCAAGCAGCACTGGAGAC 1042  
QY 3273 CTGCATATCTGGCTTCTACGGTATCCCAACAAATGGAGGAAATGTGACCCATCAAGTG 3332  
Db 1043 CTGCATATCTGGCTTCTACGGTATCCCAACAAATGGAGGAAATGTGACCCATCAAGTG 1102  
QY 3333 CAATGGGACGGCTCTGTGCAACCAACCAACGGGCAAGTCTTCGACCAACCAAGG 3392  
Db 1103 CAATGGGACGGCTCTGTGCAACCAACCAACGGGCAAGTCTTCGACCAACCAAGG 1162  
QY 3393 COTCAAGGGGACGAGTGCCAGCTATGTGAGTGAATAATCGATACCAAGAAACCCCTCT 3452  
Db 1163 COTCAAGGGGACGAGTGCCAGCTATGTGAGTGAATAATCGATACCAAGAAACCCCTCT 1222  
QY 3453 CAGAGGAACATGTTATATATATCTCTTATTGACTATCAGTTACCTTTAGTCTATCCCA 3512  
Db 1223 CAGAGGAACATGTTATATATATCTCTTATTGACTATCAGTTACCTTTAGTCTATCCCA 1282  
QY 3513 GGAAGATGATCGCTATTACACAGCTATCAATTTTGTGCTACTCTGACGAAACAAACAG 3572  
Db 1283 GGAAGATGATCGCTATTACACAGCTATCAATTTTGTGCTACTCTGACGAAACAAACAG 1342  
QY 3573 GGAATTTGACATGTTCAATATGCTTCAAGAAATTCACCTCAACATCAGCTGGGTGC 3632  
Db 1343 GGAATTTGACATGTTCAATATGCTTCAAGAAATTCACCTCAACATCAGCTGGGTGC 1402  
QY 3633 CAGTTTCTCAGCTGAACCCAGGCTGGAGAGAGATGCTGTCTTTCAAAACCAACAT 3692  
Db 1403 CAGTTTCTCAGCTGAACCCAGGCTGGAGAGAGATGCTGTCTTTCAAAACCAACAT 1462  
QY 3693 TAAGGAGTACAAAGATGTTCTTCTAATGAGAAGTTGATTTTCGAAACCAACCAATAT 3752  
Db 1463 TAAGGAGTACAAAGATGTTCTTCTAATGAGAAGTTGATTTTCGAAACCAACCAATAT 1522  
QY 3753 CACTTTCTTTGTTATGTCAGTAATTCACCTGGCCCATCAAAATTCAGATGCTCTC 3812  
Db 1523 CACTTTCTTTGTTATGTCAGTAATTCACCTGGCCCATCAAAATTCAGATGCTCTC 1582  
QY 3813 TCAGCACAGCAATTTATGACCTGGTACAGTCTTCTGTCAGTCTTCTTCTTCTTCT 3872  
Db 1583 TCAGCACAGCAATTTATGACCTGGTACAGTCTTCTGTCAGTCTTCTTCTTCTTCT 1642  
QY 3873 CTCCTTGTCTCTGGTCTGTGTTTGAAGATCAAAACAAAGTTGTTGGGCTCCAG 3932  
Db 1643 CTCCTTGTCTCTGGTCTGTGTTTGAAGATCAAAACAAAGTTGTTGGGCTCCAG 1702  
QY 3933 ACGTAGAGAGCAACTTCTTCAGAGATGCAACAGATGCGCAGCGCTCCCTTGTCTCTGT 3992  
Db 1703 ACGTAGAGAGCAACTTCTTCAGAGATGCAACAGATGCGCAGCGCTCCCTTGTCTCTGT 1762  
QY 3993 AAATGTCCCTTGGAAACAGATGAGGAGCTCTCTGATCTTATTGGGGGAGTATAAGAC 4052  
Db 1763 AAATGTCCCTTGGAAACAGATGAGGAGCTCTCTGATCTTATTGGGGGAGTATAAGAC 1822  
QY 4053 TGTTCCTCAACCCATTGCACTGGAGCGGTGTTTGGCAACAAAGCCCTCTCTCTCTGT 4112  
Db 1823 TGTTCCTCAACCCATTGCACTGGAGCGGTGTTTGGCAACAAAGCCCTCTCTCTCTGT 1882  
QY 4113 GTTTGTGAGGCTCCTCGAGCCCTGGGTGGCATCCCTCTCTGGGAGTCAAGTCTTGC 4172  
Db 1883 GTTTGTGAGGCTCCTCGAGCCCTGGGTGGCATCCCTCTCTGGGAGTCAAGTCTTGC 1942  
QY 4173 TGTGGCAGCGCTCTGTGGATCTTCTCAGCAGATGCGCATAGTGTACAGGAGAGTC 4232  
Db 1943 TGTGGCAGCGCTCTGTGGATCTTCTCAGCAGATGCGCATAGTGTACAGGAGAGTC 2002  
QY 4233 AGGAGCCGTGAGAAACCGGAGCAGCAGCCCTCTGCAAGCCCTGGGACCTGCACTGCA 4290  
Db 2003 AGGAGCCGTGAGAAACCGGAGCAGCAGCCCTCTGCAAGCCCTGGGACCTGCACTGCA 2060

AAS72660  
ID AAS72660 standard; cDNA; 3490 BP.  
XX  
AC AAS72660;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #8464.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR  
PR 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
DR  
DR P-PSDB; ABG08473.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 1; SEQ ID NO 8464; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (II) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 3490 BP; 931 A; 772 C; 858 G; 929 T; 0 U; 0 Other;

Query Match 25.1%; Score 1077; DB 5; Length 3490;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1597; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 409 AGACTAACTGGATCTTCTGGTGTGTGACATGGACCTGGAAATTTAATAACAACG 468

Db 22 AGACTAACTGGATCTTCTGGTGTGTGACATGGACCTGGAAATTTAATAACAACG 81

QY 469 AAGTCCAGTGGCTCATTTGAAGGACGACCAATAGAATTAATGAGACTTCGTTTCAATCAT 528

Db 82 AAGTCCAGTGGCTCATTTGAAGGACGACCAATAGAATTAATGAGACTTCGTTTCAATCAT 141



CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activities. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 3165 BP; 1042 A; 705 C; 687 G; 731 T; 0 U; 0 Other;

Query Match 10.0%; Score 428; DB 5; Length 3165;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-189;  
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2757 AAGGCTGCAACACAGTGTAGCAGTGCAGACACCATGTGCTTGAGCAGCATG 2816  
 DB |||||  
 QY 1218 AAGGCTGCAACACAGTGTAGCAGTGCAGACACCATGTGCTTGAGCAGCATG 1277  
 DB |||||  
 QY 2817 TGGAGATTGCACACGCGCAGCTGTAGTGTGTCAGCAACATGAAGCAGTGTGT 2876  
 DB |||||  
 QY 1278 TGGAGATTGCACACGCGCAGCTGTAGTGTGTCAGCAACATGAAGCAGTGTGT 1337  
 DB |||||  
 QY 2877 GGACTCCAAATGCTATGGCTCTTCCCTTTGGCCAGTGTATGGAATGGTATACGAT 2936  
 DB |||||  
 QY 1338 GGACTCCAAATGCTATGGCTCTTCCCTTTGGCCAGTGTATGGAATGGTATACGAT 1397  
 DB |||||  
 QY 2937 GAGCAGCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAATGCTGGAGCA 2996  
 DB |||||  
 QY 1398 GAGCAGCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAATGCTGGAGCA 1457  
 DB |||||  
 QY 2997 ACCAGGCTGTGGCTGTACTGTATCCAGCAATACTGGCAAGGGAATGCTATAGAGG 3056  
 DB |||||  
 QY 1458 ACCAGGCTGTGGCTGTACTGTATCCAGCAATACTGGCAAGGGAATGCTATAGAGG 1517  
 DB |||||  
 QY 3057 TTCCTATAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTATCCACA 3116  
 DB |||||  
 QY 1518 TTCCTATAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTATCCACA 1577  
 DB |||||  
 QY 3117 GCGCTGTCAATTCAGCATGTGCTATAGAGACAGCAGATACAACTGCTTTCATTCA 1637  
 DB |||||

## RESULT 14

AAS76294  
 ID AAS76294 standard; cDNA; 3313 BP.

XX  
 AC AAS76294;

XX  
 DT 13-FEB-2002 (first entry)

XX  
 DE DNA encoding novel human diagnostic protein #12098.

XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX  
 OS Homo sapiens.

XX  
 FN WO200175067-A2.

XX  
 PD 11-OCT-2001.

XX

PF 30-MAR-2001; 2001WO-US008631.  
 XX  
 PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PA  
 XX Drmanac RT, Liu C, Tang YT;  
 PI  
 XX  
 DR WPI; 2001-639362/73.  
 DR P-PSDB; ABG12107.  
 XX  
 PT New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 1; SEQ ID NO 12098; 103pp; English.  
 XX  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activities. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 3313 BP; 1050 A; 731 C; 724 G; 768 T; 0 U; 0 Other;

Query Match 10.0%; Score 428; DB 5; Length 3313;  
 Best Local Similarity 100.0%; Pred. No. 2.7e-189;  
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2757 AAGGCTGCAACACAGTGTAGCAGTGCAGACACCATGTGCTTGAGCAGCATG 2816  
 DB |||||  
 DB 1218 AAGGCTGCAACACAGTGTAGCAGTGCAGACACCATGTGCTTGAGCAGCATG 1277  
 DB |||||  
 QY 2817 TGGAGATTGCACACGCGCAGCTGTAGTGTGTCAGCAACATGAAGCAGTGTGT 2876  
 DB |||||  
 DB 1278 TGGAGATTGCACACGCGCAGCTGTAGTGTGTCAGCAACATGAAGCAGTGTGT 1337  
 DB |||||  
 QY 2877 GGACTCCAAATGCTATGGCTCTTCCCTTTGGCCAGTGTATGGAATGGTATACGAT 2936  
 DB |||||  
 DB 1338 GGACTCCAAATGCTATGGCTCTTCCCTTTGGCCAGTGTATGGAATGGTATACGAT 1397  
 DB |||||  
 QY 2937 GAGCAGCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAATGCTGGAGCA 2996  
 DB |||||  
 DB 1398 GAGCAGCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAATGCTGGAGCA 1457  
 DB |||||  
 QY 2997 ACCAGGCTGTGGCTGTACTGTATCCAGCAATACTGGCAAGGGAATGCTATAGAGG 3056  
 DB |||||  
 DB 1458 ACCAGGCTGTGGCTGTACTGTATCCAGCAATACTGGCAAGGGAATGCTATAGAGG 1517  
 DB |||||  
 QY 3057 TTCCTATAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTATCCACA 3116  
 DB |||||  
 DB 1518 TTCCTATAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTATCCACA 1577  
 DB |||||  
 QY 3117 GCGCTGTCAATTCAGCATGTGCTATAGAGACAGCAGATACAACTGCTTTCATTCA 3176  
 DB |||||

Db 1578 GCCCTGCTCAATCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCAATCA 1637

QY 3177 CTGTCCAG 3184

Db 1638 CTGTCCAG 1645

RESULT 15

AA02409

ID AAS02409 standard; cDNA; 668 BP.

XX AC AAS02409;

XX DT 18-JUL-2001 (first entry)

XX DE Human secreted protein, cDNA #15.

XX KW Human; secreted protein; immunogen; antibody; diagnosis;

XX KW rheumatoid arthritis; hyperproliferative disorder; neoplasm;

XX KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;

XX KW angiogenesis; Alzheimer's disease; bacterial infection; viral infection;

XX KW fungal infection; corneal infection; wound healing; cell culture;

XX KW epithelial cell proliferation; skin ageing; transplantation;

XX KW tissue regeneration; chemotaxis; food additive; ss.

XX OS Homo sapiens.

XX PN WO200123546-A1.

XX PD 05-APR-2001.

XX PF 26-SEP-2000; 2000WO-US026323.

XX PR 27-SEP-1999; 99US-0155805P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben S, Komatsoulis GA;

XX DR WPI; 2001-266150/27.

XX DR P-PSDB; AAU01736, AAU01774.

XX PT Nucleic acids encoding 37 human secreted polypeptides, useful for

XX PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease

XX PT and diabetic retinopathy.

XX PS Disclosure; Page 420; 494pp; English.

XX CC The sequence encodes a human secreted protein of the invention. The

XX CC polynucleotides, polypeptides and antibodies raised against them are used

XX CC to prevent, treat or ameliorate a medical condition in e.g. humans, mice,

XX CC rabbits, goats, horses, cats, dogs, chickens or sheep. The

XX CC polynucleotides and antibodies are also used in diagnosing a pathological

XX CC condition or susceptibility to a pathological condition. The antibodies

XX CC can also be used in alleviating symptoms associated with the disorders

XX CC and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked

XX CC immunosorbent assays (ELISA). Disorders which are diagnosed or treated

XX CC include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative

XX CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders

XX CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,

XX CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,

XX CC infections caused by bacteria, viruses and fungi and ocular disorders

XX CC e.g. corneal infection. The polypeptides can also be used to aid wound

XX CC healing and epithelial cell proliferation, to prevent skin aging due to

XX CC sunburn, to maintain organs before transplantation, for supporting cell

XX CC culture of primary tissues, to regenerate tissues and in chemotaxis. The

XX CC polypeptides can also be used as a food additive or preservative to

XX CC increase or decrease storage capabilities

XX SQ Sequence 668 BP; 298 A; 113 C; 120 G; 137 T; 0 U; 0 Other;

Query Match 7.3%; Score 314; DB 4; Length 668;

Best Local Similarity 100.0%; Pred. No. 5.5e-136;

Matches 314; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3488 ATCAGTTACCTTTAGTCTATCCAGGAAGATGATCGTATTACAGCTATCAATTTG 3547

Db 9 ATCAGTTACCTTTAGTCTATCCAGGAAGATGATCGTATTACAGCTATCAATTTG 68

QY 3548 TGGCTACTCTGACGAACAACACAGGGATTGGACATGTTTCATCATGCTCTCAAGAATT 3607

Db 69 TGGCTACTCTGACGAACAACACAGGGATTGGACATGTTTCATCATGCTCTCAAGAATT 128

QY 3608 TCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGA 3667

Db 129 TCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGA 188

QY 3668 TGGCTGTGTTTCAAAAAACAACATTAAAGAGTACAAAGATGTTTCTTAATGAGAAGT 3727

Db 189 TGGCTGTGTTTCAAAAAACAACATTAAAGAGTACAAAGATGTTTCTTAATGAGAAGT 248

QY 3728 TTCAATTTTGGCAACCCCAAAATATCACTTTCTTTGTTTATGTCAGTAATTTTCACTGGC 3787

Db 249 TTGATTTTGGCAACCCCAAAATATCACTTTCTTTGTTTATGTCAGTAATTTTCACTGGC 308

QY 3788 CCATCAAAATTCAG 3801

Db 309 CCATCAAAATTCAG 322

RESULT 16

AAS72658

ID AAS72658 standard; cDNA; 671 BP.

XX AC AAS72658;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #8462.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG08471.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

XX PT diagnostics, forensics, gene mapping, identification of mutations

XX PT responsible for genetic disorders or other traits and to assess

XX PT biodiversity.

XX PS Claim 1; SEQ ID NO 8462; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)

XX CC sequences. (I) is useful as hybridisation probes, polymerase chain

XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

XX CC and in recombinant production of (II). The polynucleotides are also used

XX CC in diagnostics as expressed sequence tags for identifying expressed

XX CC genes. (I) is useful in gene therapy techniques to restore normal

XX CC activity of (II) or to treat disease states involving (II). (II) is

XX CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (ii) and its binding partners are useful in medical imaging  
 CC of sites expressing (ii). (i) and (ii) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostic, forensic, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 671 BP; 207 A; 147 C; 148 G; 169 T; 0 U; 0 Other;  
 Query Match 7.0%; Score 300; DB 5; Length 671;  
 Best Local Similarity 100.0%; Pred. No. 2e-129;  
 Matches 300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3502 AGTCTATCCCGAAGATGATCGCTATTACACAGCTATCAATTTTGGCTACTCTGAC 3561  
 Db |||||  
 QY 155 AGTCTATCCCGAAGATGATCGCTATTACACAGCTATCAATTTTGGCTACTCTGAC 214  
 Db |||||  
 QY 3562 GAACAAAACAGGATTTGGACATGTTCAATGCTCCAAAGATTTCAACCTCAACATC 3621  
 Db |||||  
 QY 215 GAACAAAACAGGATTTGGACATGTTCAATGCTCCAAAGATTTCAACCTCAACATC 274  
 Db |||||  
 QY 3622 ACCTGGCTGCGAGTTTCTACGTGGAACCCAGGCTGGAGAGATGCTGTTGTTTCA 3691  
 Db |||||  
 QY 275 ACCTGGCTGCGAGTTTCTACGTGGAACCCAGGCTGGAGAGATGCTGTTGTTTCA 334  
 Db |||||  
 QY 3692 AAACCAACATTAAGGATGACAAAGATAGTTTCTATGATGAGAGTTTGATTTTGGCAAC 3741  
 Db |||||  
 QY 335 AAACCAACATTAAGGATGACAAAGATAGTTTCTATGATGAGAGTTTGATTTTGGCAAC 394  
 Db |||||  
 QY 3742 CACCCAAATATCACATTTCTTTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 3801  
 Db 395 CACCCAAATATCACATTTCTTTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAG 454  
 RESULT 17  
 ID AAZ91916  
 AC AAZ91916;  
 XX  
 DT 08-JUN-2000 (first entry)  
 DE Wild type (C57BL/6J) murine mahogany protein coding sequence.  
 KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;  
 KW weight regulation; cell therapy; body weight disorder; cachexia;  
 KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
 KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.  
 XX  
 OS Mus sp.  
 PN WO200005373-A2.  
 XX  
 PD 03-FEB-2000.  
 XX  
 PF 21-JUL-1999; 99WO-US016484.  
 XX  
 PR 21-JUL-1998; 98US-0093630P.  
 PR 20-OCT-1998; 98US-0104978P.  
 PR 05-FEB-1999; 99US-00245041.  
 XX  
 PA (MILL-) MILLENIUM PHARM INC.  
 XX  
 PI Moore K, Nagle DL;  
 XX  
 DR WPI; 2000-195103/17.

DR P-PSDB; AAY81803.  
 XX New human and murine mahogany genes, useful, e.g. for diagnosis and  
 PT treatment of body weight disorders.  
 XX  
 PS Claim 1; Fig 2a; 188pp; English.  
 XX  
 CC This sequence represents a murine mahogany gene of the invention. The  
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (ii); (ii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iii) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in (cell-  
 CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (ii) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (ii) (for diagnosis, prognosis or monitoring of treatment); to  
 CC evaluate (ii)-expressing cells intended for cell therapy, and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity. (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other  
 CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate  
 XX  
 SQ Sequence 8827 BP; 2304 A; 2077 C; 2162 G; 2284 T; 0 U; 0 Other;  
 Query Match 6.8%; Score 291; DB 3; Length 8827;  
 Best Local Similarity 100.0%; Pred. No. 3.1e-125;  
 Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4000 GCCTTGGAAACAGATGAGGAGCTCTCTGATCTTATTTGGGGGAGTATAAGACTGTTC 4059  
 Db 4266 GCCTTGGAAACAGATGAGGAGCTCTCTGATCTTATTTGGGGGAGTATAAGACTGTTC 4325  
 QY 4060 AAACCCATTTGCACTGAGGCGCTGTTTGGCAACAAAGCGCTGCTCTCTCTGTTGTG 4119  
 Db 4326 AAACCCATTTGCACTGAGGCGCTGTTTGGCAACAAAGCGCTGCTCTCTCTGTTGTG 4385  
 QY 4120 AGCTCCCTCGAGGCGCTGGTGGCATCCCTCTCTGGSCAGTCAAGTCTTCTGTGGCC 4179  
 Db 4386 AGCTCCCTCGAGGCGCTGGTGGCATCCCTCTCTGGSCAGTCAAGTCTTCTGTGGCC 4445  
 QY 4180 AGCGCCCTGCTGGACATTTCTCAGCAGATGCGCATAGTGTACAAGAGAAGTCAGAGCC 4239  
 Db 4446 AGCGCCCTGCTGGACATTTCTCAGCAGATGCGCATAGTGTACAAGAGAAGTCAGAGCC 4505  
 QY 4240 GTGAGAAACCGAAGCAGCAGCCCTGCAACAGCTGGGACCTGCATCTGA 4290  
 Db 4506 GTGAGAAACCGAAGCAGCAGCCCTGCAACAGCTGGGACCTGCATCTGA 4556  
 RESULT 18  
 ID ABK38928/c  
 AC ABK38928 standard; cDNA; 246 BP.  
 XX  
 DT 21-MAY-2002 (first entry)  
 DE cDNA encoding lung tumour protein clone R0127:F03.  
 KW Lung tumour; cancer; T cell; immune response stimulator; cytostatic;  
 KW gene; ss.  
 XX  
 OS Homo sapiens.  
 PN WO200204514-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 DR 10-JUL-2001; 2001WO-US022058.

```
XX PR 11-JUL-2000; 2000US-00614124.
XX PR 29-AUG-2000; 2000US-00651563.
XX PR 08-SEP-2000; 2000US-00658824.
XX PR 26-SEP-2000; 2000US-00671325.
XX PR 06-OCT-2000; 2000US-00677419.
XX PR 30-OCT-2000; 2000US-00702705.
XX PR 13-DEC-2000; 2000US-00736457.
XX PR 03-MAY-2001; 2001US-00849626.
XX PA (CORI-) CORIYA CORP.
XX XX
XX PI Wang T, Watanabe Y, Henderson RA, Johnson JC, Retter MW;
XX PI Marxerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS, Mcnabb A;
XX PI Wang A, Fanger N, Switzer A, Mcneill PD, Clapper JD;
XX DR WPI; 2002-164634/21.
XX XX
XX PT Novel polynucleotide encoding a lung tumor polypeptide useful for
XX PT stimulating and/or expanding T cells specific for a tumor protein.
XX XX
XX PS Claim 1; SEQ ID NO 966; 223pp; English.
XX XX
XX CC The invention describes an isolated polynucleotide and polypeptide useful
XX CC for stimulating and/or expanding T cells specific for a tumour protein
XX CC containing the presence of a cancer in a patient. A composition
XX CC containing the polynucleotide and/or polypeptide is useful for treating a
XX CC lung cancer in a patient. The polypeptide is useful for removing tumour
XX CC cells from a biological sample. The polynucleotide is also useful as
XX CC probe or primer to detect the level of mRNA encoding a tumour protein.
XX CC This sequence encodes a lung tumour associated protein or protein
XX CC fragment, described in the method of the invention. Note: The sequence
XX CC data for this patent did not form part of the printed specification, but
XX CC was obtained in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX XX
XX SQ Sequence 246 BP; 63 A; 61 C; 60 G; 62 T; 0 U; 0 Other;
XX XX
XX Query Match 5.7%; Score 246; DB 6; Length 246;
XX Best Local Similarity 100.0%; Freq. No. 3.5e-104;
XX Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX QY 2515 CTGGAATAATGTCAGTCTCTCAGAGCATGTCACAGCTCACCTTAACCCATGGTGGC 2574
XX DB 246 CTGCGAATAATGTCAGTCTCTCAGAGCATGTCACAGCTCACCTTAACCCATGGTGGC 187
XX QY 2575 CTTCCGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCCATTTACAAATAGT 2634
XX DB 186 CTTCCGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCCATTTACAAATAGT 127
XX QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTCTGATCTGGATCTGTGGAATTTATCAGAA 2694
XX DB 126 TTACTACAGTGGATCCGCTCTGAGCCAGTCTGATCTGGATCTGTGGAATTTATCAGAA 67
XX QY 2695 CCCAGTACTCGGGACTGAAGGCTGCACCTGTCATCAACCCACTCAATGGTGTCTGT 2754
XX DB 66 CCCAGTACTCGGGACTGAAGGCTGCACCTGTCATCAACCCACTCAATGGTGTCTGT 7
XX QY 2755 GAAAGG 2760
XX DB 6 GAAAGG 1
XX XX
XX RESULT 19
XX ACAL1257/c
XX ID ACAL1257 standard; cDNA; 246 BP.
XX AC ACAL1257;
XX XX
XX DT 05-JUN-2003 (first entry)
XX DE Human lung adenocarcinoma library cDNA SEQ ID 966.
XX XX
```

```
KW Human; lung cancer; ss; lung tumour; cytostatic; vaccine;
KW T cell expansion; CD4; CD8.
XX OS Homo sapiens.
XX PN US2002197669-A1.
XX PD 26-DEC-2002.
XX PF 03-MAY-2001; 2001US-00849626.
XX PR 13-DEC-2000; 2000US-00736457.
XX XX
XX PA (BANG/) BANGUR C S.
XX PA (FANG/) FANGER G R.
XX PA (WANG/) WANG A.
XX PA (WANG/) WANG T.
XX PA (SWIT/) SWITZER A P.
XX PA (MCNE/) MCNEILL P D.
XX PA (CLAP/) CLAPPER J D.
XX XX
XX PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, Mcneill PD;
XX PI Clapper JD;
XX XX
XX WPI; 2003-352750/33.
XX XX
XX PT Novel lung cancer polynucleotide encoding lung cancer protein, useful for
XX PT detecting the presence of lung cancer in a patient, and in pharmaceutical
XX PT compositions, e.g. vaccines, for treating lung cancer.
XX XX
XX PS Example 1; Page; 72pp; English.
XX XX
XX CC The invention relates to a polynucleotide encoding a lung tumour protein,
XX CC comprising a sequence selected from any of the 14 sequences mentioned in
XX CC the specification, or a sequence (S2) mentioned in specification,
XX CC complement of S1, sequences consisting of at least 20 contiguous residues
XX CC of S1, sequences that hybridise to S1, sequences having 75%, preferably
XX CC 90%, identity to S1, or degenerate variants of S1. Also included are an
XX CC isolated polypeptide (comprising a sequence (S3) selected from any one of
XX CC the 4 amino acid sequences mentioned in the specification, a sequence
XX CC encoded by the polynucleotide, or sequences having at least 70%,
XX CC preferably 90%, identity to a sequence encoded by the polynucleotide), an
XX CC expression vector comprising the polynucleotide operably linked to an
XX CC expression control sequence, a host cell transformed or transfected with
XX CC the vector, an isolated antibody (or its antigen-binding fragment) that
XX CC specifically binds to the polypeptide, detecting the presence of a cancer
XX CC in a patient, a fusion protein comprising the polypeptide, an
XX CC oligonucleotide that hybridises to S1 under moderately stringent
XX CC conditions, stimulating and/or expanding T cells specific for a tumour
XX CC protein (comprising contacting T cells with the polynucleotide, protein
XX CC or antigen-presenting cells, under conditions and for a time sufficient
XX CC to permit the stimulation and/or expansion of T cells) and inhibiting the
XX CC development of a cancer in a patient (by incubating CD4+ and/or CD8+ T
XX CC cells isolated from a patient with the polynucleotide, protein or antigen
XX CC presenting cells that express the polynucleotide, such that T cells
XX CC proliferate, administering to the patient an effective amount of the
XX CC proliferated T cells, and thus inhibiting the development of a cancer in
XX CC the patient. The polynucleotide, protein and cells are useful in a
XX CC composition for stimulating an immune response in a patient, and for
XX CC treating a cancer in a patient (particularly lung cancer). The
XX CC oligonucleotide is useful for determining the presence of a cancer in a
XX CC patient. The protein and oligonucleotides are useful in pharmaceutical
XX CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe
XX CC or primer for nucleic acid hybridisation, and in the design and
XX CC preparation of ribozyme molecules for inhibiting expression of tumour
XX CC polypeptides and proteins in tumour cells. An amplified portion of the
XX CC polynucleotide is useful for isolating a full-length gene from a suitable
XX CC library. The present sequence is a cDNA (full length, extended or
XX CC partial) isolated from a library derived from lung tumour/cancer cells.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from the
XX CC USPTO at seqdata.uspto.gov/sequence.html?DocId=20020197669
```

```
SQ Sequence 246 BP; 63 A; 61 C; 60 G; 62 T; 0 U; 0 Other;
Query Match 5.7%; Score 246; DB 7; Length 246;
Best Local Similarity 100.0%; Pred. No. 3.5e-104;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2515 CTGCGAATAATGCAATGTCCTTACTGCTGGGAGATATGTCCTCCCAATTTACAATAAGT 2634
Db 246 CTGCGAATAATGCAATGTCCTTACTGCTGGGAGATATGTCCTCCCAATTTACAATAAGT 127
QY 2575 CTTCGGAAGATCAATGTCCTTACTGCTGGGAGATATGTCCTCCCAATTTACAATAAGT 2634
Db 186 CTTCGGAAGATCAATGTCCTTACTGCTGGGAGATATGTCCTCCCAATTTACAATAAGT 127
QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTATGTCCTGGAATTTATCAGAA 2694
Db 126 TTACTACAGTGGATCCGCTCTGAGCCAGTATGTCCTGGAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGGAAGCTGCAACCTGCAATCAACCCCACTCAATGGTAGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGGAAGCTGCAACCTGCAATCAACCCCACTCAATGGTAGTCTGT 7
QY 2755 GAAAGG 2760
Db 6 GAAAGG 1
RESULT 20
ACAO2443/C
ID ACA02443 standard; cDNA; 246 BP.
AC ACA02443;
DT 22-MAY-2003 (first entry)
DE Lung cancer therapy and diagnosis associated cDNA #932.
DE Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ss.
OS Homo sapiens.
PN US2002172952-A1.
PD 21-NOV-2002.
PF 10-JUL-2001; 2001US-00902941.
PR 30-JUN-1999; 99US-00346492.
PR 15-OCT-1999; 99US-00419356.
PR 17-DEC-1999; 99US-00466887.
PR 30-DEC-1999; 99US-00476300.
PR 06-MAR-2000; 2000US-00519642.
PR 12-MAR-2000; 2000US-00533077.
PR 27-APR-2000; 2000US-00546259.
PR 05-JUN-2000; 2000US-00589184.
PR 11-JUL-2000; 2000US-00614124.
PR 29-AUG-2000; 2000US-00651563.
PR 08-SEP-2000; 2000US-00658824.
PR 26-SEP-2000; 2000US-00671325.
PR 06-OCT-2000; 2000US-00677419.
PR 30-OCT-2000; 2000US-00702705.
PR 13-DEC-2000; 2000US-00736457.
PR 03-MAY-2001; 2001US-00849626.
XX (CORI-) CORIXA CORP.
XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW;
PI Durham M, Carter D, Fanger GR, Vedvick TS, Bangur CS, McNabb A;
XX WPI; 2003-328427/31.
XX New polynucleotide, useful for preparing a composition for treating or
PT inhibiting development of cancer, e.g. lung cancer.
XX Example 1; SEQ ID NO 966; 82pp; English.
XX The invention describes an isolated polynucleotide comprising one of 32
CC sequences, complement or degenerate variants of them. The polynucleotide
CC is useful for preparing a composition e.g. a vaccine or for gene therapy,
CC for treating or inhibiting development of cancer, e.g. lung cancer. This
CC sequence represents a polynucleotide associated with the compositions and
CC methods for the therapy and diagnosis of lung cancer
XX
SQ Sequence 246 BP; 63 A; 61 C; 60 G; 62 T; 0 U; 0 Other;
Query Match 5.7%; Score 246; DB 7; Length 246;
Best Local Similarity 100.0%; Pred. No. 3.5e-104;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2515 CTGCGAATAATGCAATGTCCTTACTGCTGGGAGATATGTCCTCCCAATTTACAATAAGT 2574
Db 246 CTGCGAATAATGCAATGTCCTTACTGCTGGGAGATATGTCCTCCCAATTTACAATAAGT 187
QY 2575 CTTCGGAAGATCAATGTCCTTACTGCTGGGAGATATGTCCTCCCAATTTACAATAAGT 2634
Db 186 CTTCGGAAGATCAATGTCCTTACTGCTGGGAGATATGTCCTCCCAATTTACAATAAGT 127
QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTATGTCCTGGAATTTATCAGAA 2694
Db 126 TTACTACAGTGGATCCGCTCTGAGCCAGTATGTCCTGGAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGGAAGCTGCAACCTGCAATCAACCCCACTCAATGGTAGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGGAAGCTGCAACCTGCAATCAACCCCACTCAATGGTAGTCTGT 7
QY 2755 GAAAGG 2760
Db 6 GAAAGG 1
RESULT 21
ABZ72040
ID ABZ72040 standard; DNA; 207433 BP.
XX AC ABZ72040;
XX DT 03-APR-2003 (first entry)
XX DE Gene 216 H194BAC1098L22 nucleotide sequence SEQ ID NO 5.
XX Human; Gene 216; chromosome 20p13-p12; antiasthmatic; anorectic;
XX antiinflammatory; gastrointestinal; gene therapy; vaccine; asthma;
XX obesity; inflammatory bowel disease; promoter; gene; ss.
XX Homo sapiens.
XX WO200178894-A2.
XX 25-OCT-2001.
XX 13-APR-2001; 2001WO-US012245.
XX 13-APR-2000; 2000US-00548797.
XX (GENO-) GENOME THERAPEUTICS CORP.
XX Keith T;
XX WPI; 2001-639428/73.
XX P-PSDB; ABR00926.
XX Isolated genes (Gene 216) from human chromosome 20p13-p12 and the
XX proteins they encode, useful for the prevention, diagnosis and treatment
XX of asthma, obesity and inflammatory bowel disease.
XX
```



PS Example 4; Fig 7; 520pp; English.

XX The invention relates to isolated genes (Gene 216) from human chromosome

CC 20p13-p12 and the proteins they encode. The nucleic acids and proteins

CC may be used in the prevention, diagnosis and treatment of diseases

CC associated with inappropriate Gene 216 expression. For example, the

CC nucleic acids (or vectors) and proteins may be used to treat disorders

CC associated with decreased expression by rectifying mutations or deletions

CC in a patient's genome that affect the activity of gene 216 by expressing

CC inactive proteins or to supplement the patients own production of Gene

CC 216 proteins. Additionally, the nucleic acids may be used to produce the

CC secreted Gene 216 protein, by inserting the nucleic acids into a host

CC cell and culturing the cell to express the protein. The nucleic acids and

CC complementary sequences may also be used as DNA probes in diagnostic

CC assays to detect and quantitate the presence of similar nucleic acid

CC sequences in samples and therefore which patients may be in need of

CC restorative therapy. The Gene 216 protein may also be used as antigens in

CC the production of antibodies against Gene 216 and in assays to identify

CC modulators of Gene 216 expression and activity. The anti-Gene 216

CC antibodies and antagonists may also be used to down regulate expression

CC and activity. The anti-Gene 216 antibodies may also be used as diagnostic

CC agents for detecting the presence of Gene 216 proteins in samples (e.g.

CC by enzyme linked immunosorbant assay or ELISA). Disorders that may be

CC prevented, diagnosed and/or treated by the above methods include, for

CC example asthma, obesity and inflammatory bowel disease. The present

CC sequence is that of the Gene 216 genomic nucleic acid sequence, promoter

CC or enhancer

XX

SQ Sequence 207433 BP; 52775 A; 51290 C; 51698 G; 51670 T; 0 U; 0 Other;

Query Match 5.5%; Score 234; DB 5; Length 207433;

Best Local Similarity 100.0%; Pred. No. 1.4e-98;

Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2951 CTGAAAATTGTTCCAGGCTACTGTACCTGTAGTCTGTGAGCAACACCGCTGTGGCT 3010

DB 9813 CTGAAAATTGTTCCAGGCTACTGTACCTGTAGTCTGTGAGCAACACCGCTGTGGCT 9872

QY 3011 GGTGTACTGATCCAGCAATCTGCAAGGGAATGCATAGAGGTTCTCTATAAAGGAC 3070

DB 9873 GGTGTACTGATCCAGCAATCTGCAAGGGAATGCATAGAGGTTCTCTATAAAGGAC 9932

QY 3071 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATT 3130

DB 9933 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATT 9992

QY 3131 CCAGCATGTGCTAGAGGACAGCATACACTGGTCTTTTCATTCTCTCTGCTCCAG 3184

DB 9993 CCAGCATGTGCTAGAGGACAGCATACACTGGTCTTTTCATTCTCTCTGCTCCAG 10046

RESULT 22

ABX74891

ID ABX74891 standard; DNA; 207433 BP.

XX

AC ABX74891;

XX

DT 07-APR-2003 (first entry)

XX

DE BAC1098L22 DNA sequence.

XX

KW Gene 216; antiasthmatic; antiinflammatory; ss; anorectic;

KW chromosome 20p13-p12; single nucleotide polymorphism; SNP; gene therapy;

KW respiratory disease; asthma; obesity; bronchial hyper-responsiveness;

KW chronic obstructive pulmonary disease;

KW adult respiratory distress syndrome; inflammatory bowel syndrome.

XX

OS Synthetic.

XX

PN WC0200283077-A2.

XX

XX 24-OCT-2002.

PD

XX

PF 15-APR-2002; 2002WO-US012063.

XX

PR 13-APR-2001; 2001US-00834597.

PR 13-APR-2001; 2001WO-US012245.

XX

PA (SCHE ) SCHERING CORP.

PA (GENO-) GENOME THERAPEUTICS CORP.

XX

PI Keith T, Little RD, Van Berdewegh P, Dupuis J, Del Mastro RG;

PI Simon J, Allen K, Pandit S;

XX

DR WPI; 2003-092960/08.

XX

PT New isolated gene 216 nucleic acids, useful for diagnosing, preventing or

PT treating a disorder, such as asthma, bronchial hyper-responsiveness,

PT chronic obstructive pulmonary disease, obesity or inflammatory bowel

PT syndrome.

XX

XX Example 6; Fig 7; 650pp; English.

PS

CC This invention relates to a novel isolated nucleic acid, gene 216,

CC identified from human chromosome 20p13-p12. The invention also discloses

CC regions of the 216 gene that contain single nucleotide polymorphisms

CC (SNP's) which may be used as markers for disease susceptibility or

CC severity. The nucleotides of the invention may have antiasthmatic,

CC antiinflammatory or anorectic activities and may be used in gene therapy.

CC The nucleic acids, antibodies or its fragments are useful for diagnosing,

CC preventing or treating a disorder, such as respiratory diseases (e.g.

CC asthma, bronchial hyper-responsiveness, chronic obstructive pulmonary

CC disease or adult respiratory distress syndrome), obesity, or inflammatory

CC bowel syndrome. The nucleic acids are also useful for identifying

CC increased susceptibility of a subject to the disorders mentioned. The

CC nucleic acids can also be used as primers and templates for the

CC recombinant production of disorder-associated peptides or polypeptides,

CC for chromosome and gene mapping, or for tissue distribution studies. The

CC present sequence represents a gene 216 cDNA sequence used in the scope of

CC the invention

XX

SQ Sequence 207433 BP; 52775 A; 51289 C; 51698 G; 51671 T; 0 U; 0 Other;

Query Match 5.5%; Score 234; DB 7; Length 207433;

Best Local Similarity 100.0%; Pred. No. 1.4e-98;

Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2951 CTGAAAATTGTTCCAGGCTACTGTACCTGTAGTCTGTGAGCAACACCGCTGTGGCT 3010

DB 9813 CTGAAAATTGTTCCAGGCTACTGTACCTGTAGTCTGTGAGCAACACCGCTGTGGCT 9872

QY 3011 GGTGTACTGATCCAGCAATCTGCAAGGGAATGCATAGAGGTTCTCTATAAAGGAC 3070

DB 9873 GGTGTACTGATCCAGCAATCTGCAAGGGAATGCATAGAGGTTCTCTATAAAGGAC 9932

QY 3071 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATT 3130

DB 9933 CAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATT 9992

QY 3131 CCAGCATGTGCTAGAGGACAGCATACACTGGTCTTTTCATTCTCTCTGCTCCAG 3184

DB 9993 CCAGCATGTGCTAGAGGACAGCATACACTGGTCTTTTCATTCTCTCTGCTCCAG 10046

RESULT 23

AAL14686/c

ID AAL14686 standard; cDNA; 481 BP.

XX

XX AAL14686;

AC

XX

DT 07-DEC-2001 (first entry)

XX

DE Human breast cancer expressed polynucleotide 7143.

XX

XX Human; breast cancer; cell marker; cytostatic; ss.

XX





CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene expression  
 CC analysis, and for identifying exons in a gene, particularly using human  
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
 CC present sequence is a single exon probe of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 531 BP; 145 A; 82 C; 107 G; 197 T; 0 U; 0 Other;

Query Match 4.9%; Score 210; DB 6; Length 531;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-87;  
 Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1238 CAACCTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATPAATGAGTCATGGTGT 1297  
 Db 184 CAACCTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATPAATGAGTCATGGTGT 243  
 QY 1298 TGTGACCCCTAAGCAAGAGCAGATGATGAGTGGTGGCACTCTGACACATTTGTTA 1357  
 Db 244 TGTGACCCCTAAGCAAGAGCAGATGATGAGTGGTGGCACTCTGACACATTTGTTA 303  
 QY 1358 CACTGAGAGTGGCGAGTGGTGCATGCTGGTGCATCTTTGGTCACTGCCCTCTCTATGGAT 1417  
 Db 304 CACTGAGAGTGGCGAGTGGTGCATGCTGGTGCATCTTTGGTCACTGCCCTCTCTATGGAT 363  
 QY 1418 ATATAAGCAATGTGACGATATGATTTGG 1447  
 Db 364 ATATAAGCAATGTGACGATATGATTTGG 393

RESULT 27  
 AA156405  
 ID AA156405 standard; DNA; 193 BP.

AC AA156405;  
 XX  
 XX  
 DT 17-OCT-2001 (first entry)  
 XX  
 XX Probe #25091 used to measure gene expression in human placenta sample.  
 DE  
 XX Probe; microarray; human; placenta; antenatal diagnosis;  
 XX Probe; microarray; ss.  
 KW  
 XX genetic disorder; ss.  
 KW  
 XX Homo sapiens.  
 OS  
 XX WO200157272-A2.  
 PN  
 XX  
 XX 09-AUG-2001.  
 PD  
 XX  
 XX 30-JAN-2001; 2001WO-US0000563.  
 PF  
 XX  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 XX (MOL3-) MOLECULAR DYNAMICS INC.  
 PA  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI  
 XX WPI; 2001-488897/53.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human placenta.  
 PT  
 XX Claim 25; SEQ ID NO 25091; 654pp; English.  
 PS  
 XX The present invention relates to single exon nucleic acid probes (SENPs).  
 XX The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders  
 XX  
 SQ Sequence 193 BP; 50 A; 36 C; 52 G; 55 T; 0 U; 0 Other;  
 Query Match 4.5%; Score 193; DB 4; Length 193;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-79;  
 Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1245 GAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTCATGGTGTGAC 1304  
 Db 1 GAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTCATGGTGTGAC 60  
 QY 1305 CCTAAGGCAAGAGGAGCAGTATGAGTGGTGGCACTCTGACACATTTGACCTGAA 1364  
 Db 61 CCTAAGGCAAGAGGAGCAGTATGAGTGGTGGCACTCTGACACATTTGACCTGAA 120  
 QY 1365 GAATGGCCGAGTGGTGCATGCTGGTGCATCTTTGGTCACTGCCCTCTCTATGATATAAG 1424  
 Db 121 GAATGGCCGAGTGGTGCATGCTGGTGCATCTTTGGTCACTGCCCTCTCTATGATATAAG 180  
 QY 1425 CAATGTGACGAA 1437  
 Db 181 CAATGTGACGAA 193

RESULT 28  
 ABS23881  
 ID ABS23881 standard; DNA; 193 BP.  
 AC ABS23881;  
 XX  
 XX 19-AUG-2002 (first entry)  
 DT  
 XX  
 XX Human genome-derived single exon probe ORF from lung SEQ ID No 23872.  
 DE  
 XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease; open reading frame; ORF.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200186003-A2.  
 EN  
 XX 15-NOV-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US0000665.  
 PF  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.

[illegible]

Db 62 GATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

QY 2725 TG 2726  
||  
Db 2 TG 1

RESULT 30  
ABA69223/c  
ID ABA69223 standard; DNA; 182 BP.  
XX AC ABA69223;  
XX DT 01-FEB-2002 (first entry)  
XX DE Human foetal liver single exon nucleic acid probe #17528.  
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
XX OS Homo sapiens.  
XX PN WO200157277-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US000669.  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX PR 26-MAY-2000; 2000US-0207456P.  
XX PR 30-JUN-2000; 2000US-00608408.  
XX PR 03-AUG-2000; 2000US-00632366.  
XX PR 21-SEP-2000; 2000US-0234687P.  
XX PR 27-SEP-2000; 2000US-0236359P.  
XX PR 04-OCT-2000; 2000GB-00024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-483447/52.  
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing  
XX PT gene expression in human fetal liver.  
XX PS Claim 4; SEQ ID NO 17528; 639pp + Sequence Listing; English.  
XX CC The invention relates to a single exon nucleic acid probe for measuring  
XX CC human gene expression in a sample derived from human foetal liver. The  
XX CC single exon nucleic acid probes may be used for predicting, measuring and  
XX CC displaying gene expression in samples derived from human fetal liver. The  
XX CC present sequence is a single exon nucleic acid probe of the invention.  
XX CC Note: The sequence data for this patent did not form part of the printed  
XX CC specification, but was obtained in electronic format directly from WIPO  
XX CC at ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;  
Query Match 4.2%; Score 182; DB 4; Length 182;  
Best Local Similarity 100.0%; Pred. No. 2.9e-74;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2545 TCCAGCTCACCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTCTCTACTGTGTC 2604  
Db 182 TCCAGCTCACCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTCTCTACTGTGTC 123

QY 2605 TGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCGTCTGAGCCCACT 2664  
Db 122 TGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCGTCTGAGCCCACT 63

QY 2665 GATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 2724  
Db 62 GATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

QY 2725 TG 2726  
||  
Db 2 TG 1

RESULT 31  
AAI49392/c  
ID AAI49392 standard; DNA; 182 BP.  
XX AC AAI49392;  
XX DT 17-OCT-2001 (first entry)  
XX DE Probe #18078 used to measure gene expression in human placenta sample.  
XX KW Probe; microarray; human; placenta; antenatal diagnosis;  
XX KW genetic disorder; ss.  
XX OS Homo sapiens.  
XX PN WO200157272-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US000663.  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX PR 26-MAY-2000; 2000US-0207456P.  
XX PR 30-JUN-2000; 2000US-00608408.  
XX PR 03-AUG-2000; 2000US-00632366.  
XX PR 21-SEP-2000; 2000US-0234687P.  
XX PR 27-SEP-2000; 2000US-0236359P.  
XX PR 04-OCT-2000; 2000GB-00024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488897/53.  
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing  
XX PT gene expression in human placenta.  
XX PS Claim 25; SEQ ID NO 18078; 654pp; English.  
XX CC The present invention relates to single exon nucleic acid probes (SENP).  
XX CC The present sequence is one such probe. The probes are useful for  
XX CC producing a microarray for predicting, measuring and displaying gene  
XX CC expression in samples derived from human placenta. The probes are useful  
XX CC for antenatal diagnosis of human genetic disorders  
XX SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;  
Query Match 4.2%; Score 182; DB 4; Length 182;  
Best Local Similarity 100.0%; Pred. No. 2.9e-74;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2545 TCCAGCTCACCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTCTCTACTGTGTC 2604  
Db 182 TCCAGCTCACCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTCTCTACTGTGTC 123

QY 2605 TGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCGTCTGAGCCCACT 2664  
Db 122 TGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCGTCTGAGCCCACT 63

QY 2665 GATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 2724  
Db 62 GATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

QY 2725 TG 2726  
||  
Db 2 TG 1

RESULT 32  
ABA51213/c  
ID ABA51213 standard; DNA; 182 BP.  
XX  
AC ABA51213;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Human breast cell single exon nucleic acid probe #9908.  
XX  
KW Human; microarray; single exon probe; gene expression; breast; disease;  
KW cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157271-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000662.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-496933/54.  
XX  
PT New spatially-addressable set of single exon nucleic acid probes, useful  
PT for measuring gene expression in sample derived from human breast,  
PT comprises number of single exon nucleic acid probes.  
XX  
PS Claim 4; SEQ ID NO 9908; 327pp + Sequence Listing; English.  
XX  
PA The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BT 474 cells. The method involves contacting the  
CC probes with a collection of detectably labelled nucleic acids derived  
CC from mRNA of human breast, and then measuring the label bound to each  
CC probe of the microarray. The probes are useful for verifying the  
CC expression of regions of genomic DNA predicted to encode proteins. They  
CC are useful for gene discovery, and for determining predisposition and/or  
CC prognosing breast disease. Gene expression analysis is useful for  
CC assessing the toxicity of chemical agents on cells. The microarray of  
CC this invention presents a far greater diversity of probes for measuring  
CC gene expression, with far less bias than expressed sequence tag  
CC microarrays. The method is suitable for rapid production of functional  
CC information from genomic sequence. The present sequence is a single exon  
CC nucleic acid probe of the invention. Note: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;  
Query Match 4.2%; Score 182; DB 4; Length 182;  
Best Local Similarity 100.0%; Pred. No. 2.9e-74;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2545 TCCAAGCTCACCTTAACCCCATGGTTCGGCCCTTCGGAAGATCAATGTCTCTACTGGTGC 2604  
DB 182 TCCAAGCTCACCTTAACCCCATGGTTCGGCCCTTCGGAAGATCAATGTCTCTACTGGTGC 123  
QY 2605 TCGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCCCTCTGAGCCCACT 2664  
DB 182 TCGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCCCTCTGAGCCCACT 123

DB 122 TCGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCCCTCTGAGCCCACT 63  
QY 2665 GATGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGAGCTGAAGGCTGCAACC 2724  
DB 62 GATGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGAGCTGAAGGCTGCAACC 3  
QY 2725 TG 2726  
DB 2 TG 1  
RESULT 33  
ABA36140/c  
ID ABA36140 standard; DNA; 182 BP.  
XX  
AC ABA36140;  
XX  
DT 23-JAN-2002 (first entry)  
XX  
DE Probe #14606 for gene expression analysis in human heart cell sample.  
XX  
KW Human; gene expression; heart; microarray; vascular system; probe;  
KW cardiovascular disease; hypertension; cardiac arrhythmia;  
KW congenital heart disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157274-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000666.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-488899/53.  
XX  
PT Single exon nucleic acid probes for analyzing gene expression in human  
PT hearts.  
XX  
PS Claim 4; SEQ ID NO 14606; 530pp; English.  
XX  
PA The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart. The  
CC present sequence is one such probe. The probes may be used for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from the human heart via microarrays. By measuring gene expression, the  
CC probes are useful for predicting, diagnosing, grading, staging,  
CC monitoring and prognosing diseases of the human heart and vascular system  
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
CC congenital heart disease. Note: The sequence data for this patent did not  
CC form part of the printed specification, but was obtained in electronic  
CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;  
Query Match 4.2%; Score 182; DB 4; Length 182;  
Best Local Similarity 100.0%; Pred. No. 2.9e-74;  
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2545 TCCAAGCTCACCTTAACCCCATGGTTCGGCCCTTCGGAAGATCAATGTCTCTACTGGTGC 2604  
DB 182 TCCAAGCTCACCTTAACCCCATGGTTCGGCCCTTCGGAAGATCAATGTCTCTACTGGTGC 123

QY 2605 TGGGAAGATATGTCCTCCATTTCACAAATAGTTTACTACAGTGGATCCGCTCTGAGCCCACT 2664  
 DB 122 TGGGAAGATATGTCCTCCATTTCACAAATAGTTTACTACAGTGGATCCGCTCTGAGCCCACT 63  
 QY 2665 GATGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGGACTGAAGCTGCAACC 2724  
 DB 62 GATGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGGACTGAAGCTGCAACC 3  
 QY 2725 TG 2726  
 DB 2 TG 1  
 RESULT 34  
 AAK43324/c  
 ID AAK43324 standard; DNA; 182 BP.  
 XX  
 AC AAK43324;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed single exon probe SEQ ID NO: 17881.  
 XX  
 DE Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 XX  
 PR 26-MAY-2000; 2000US-0207456P.  
 XX  
 PR 30-JUN-2000; 2000US-00608408.  
 XX  
 PR 03-AUG-2000; 2000US-00632366.  
 XX  
 PR 03-AUG-2000; 2000US-00632366.  
 XX  
 PR 21-SEP-2000; 2000US-0234687P.  
 XX  
 PR 27-SEP-2000; 2000US-0236359P.  
 XX  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-489300/53.  
 XX  
 DR Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human bone marrow.  
 XX  
 XX Example 4; SEQ ID NO 17881; 658pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of  
 CC the probes of the invention  
 XX  
 SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;  
 Query Match 4.2%; Score 182; DB 4; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-74;  
 Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2545 TCCAGCTCACCTTAACCCCATGGGTCCGCTTCGGAAGATCAATGTCTCTACTGTGTC 2604  
 DB 182 TCCAGCTCACCTTAACCCCATGGGTCCGCTTCGGAAGATCAATGTCTCTACTGTGTC 123  
 QY 2605 TGGGAAGATATGTCCTCCATTTCACAAATAGTTTACTACAGTGGATCCGCTCTGAGCCCACT 2664

DB 122 TGGGAAGATATGTCCTCCATTTCACAAATAGTTTACTACAGTGGATCCGCTCTGAGCCCACT 63  
 QY 2665 GATGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGGACTGAAGCTGCAACC 2724  
 DB 62 GATGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGGACTGAAGCTGCAACC 3  
 QY 2725 TG 2726  
 DB 2 TG 1  
 RESULT 35  
 AAK17517/c  
 ID AAK17517 standard; DNA; 182 BP.  
 XX  
 AC AAK17517;  
 XX  
 DT 05-NOV-2001 (first entry)  
 XX  
 DE Human brain expressed single exon probe SEQ ID NO: 17508.  
 XX  
 DE Human; brain expressed exon; gene expression analysis; probe; microarray;  
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;  
 KW ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 XX  
 PR 26-MAY-2000; 2000US-0207456P.  
 XX  
 PR 30-JUN-2000; 2000US-00608408.  
 XX  
 PR 03-AUG-2000; 2000US-00632366.  
 XX  
 PR 21-SEP-2000; 2000US-0234687P.  
 XX  
 PR 27-SEP-2000; 2000US-0236359P.  
 XX  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-483446/52.  
 XX  
 DR Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains.  
 XX  
 PS Example 4; SEQ ID NO 17508; 650pp + Sequence Listing; English.  
 XX  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is one of the probes of the  
 CC invention  
 XX  
 SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;  
 Query Match 4.2%; Score 182; DB 4; Length 182;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-74;  
 Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2545 TCCAGCTCACCTTAACCCCATGGGTCCGCTTCGGAAGATCAATGTCTCTACTGTGTC 2604  
 DB 182 TCCAGCTCACCTTAACCCCATGGGTCCGCTTCGGAAGATCAATGTCTCTACTGTGTC 123  
 QY 2605 TGGGAAGATATGTCCTCCATTTCACAAATAGTTTACTACAGTGGATCCGCTCTGAGCCCACT 2664  
 DB 122 TGGGAAGATATGTCCTCCATTTCACAAATAGTTTACTACAGTGGATCCGCTCTGAGCCCACT 63





Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2545 TCCAGCTCACCTTAACCCCATGGTGGCGCTTCGGAAGATCAATGTGCTCTACTGTGC 2604

Db 182 TCCAGCTCACCTTAACCCCATGGTGGCGCTTCGGAAGATCAATGTGCTCTACTGTGC 123

QY 2605 TGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCGGTGTGAGCCAGT 2664

Db 122 TGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCGGTGTGAGCCAGT 63

QY 2665 GATGCTGATTCGTGGGAATTTATCAGAACCCAGTACTCGGGACCTGAGGCTGCAACC 2724

Db 62 GATGCTGATTCGTGGGAATTTATCAGAACCCAGTACTCGGGACCTGAGGCTGCAACC 3

QY 2725 TG 2726

Db 2 TG 1

RESULT 39

ABS17404/c

ID ABS17404 standard; DNA; 182 BP.

XX AC ABS17404;

XX DT 19-AUG-2002 (first entry)

XX DE Human genome-derived single exon probe ORF from lung SEQ ID No 17395.

XX KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

XX KW chronic obstructive pulmonary disease; interstitial lung disease;

XX KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

XX KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

XX KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

XX KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

XX KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

XX KW primary ciliary dyskinesia; pulmonary hypertension;

XX KW hyaline membrane disease; open reading frame; ORF.

XX OS Homo sapiens.

XX PN WO200186003-A2.

XX PD 15-NOV-2001.

XX PF 30-JAN-2001; 2001WO-US000665.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024283.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2002-114183/15.

XX PT Spatially-addressable set of single exon nucleic acid probes, used to

XX PT measure gene expression in human lung samples.

XX PS Claim 4; SEQ ID NO 17395; 634pp; English.

XX CC The invention relates to a spatially-addressable set of single exon

XX CC nucleic acid probes for measuring gene expression in a sample derived

XX CC from human lung comprising single exon nucleic acid probes having one of

XX CC 12614 nucleic acid sequences mentioned in the specification, or their

XX CC complements or the 12387 open reading frames derived from the 12614

XX CC probes. Also included are a microarray comprising the novel set of probes

XX CC ; the novel set of probes which hybridise at high stringency to a nucleic

acid expressed in the human lung; measuring gene expression in a sample

derived from human lung, comprising (a) contacting the array with a

collection of detectably labeled nucleic acids derived from human lung

mRNA, and (b) measuring the label detectably bound to each probe of the

array; identifying exons in a eukaryotic genome, comprising (a)

algorithmically predicting at least one exon from genomic sequences of

the eukaryote; and (b) detecting specific hybridisation of detectably

labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

having a fragment identical to the predicted exon, the probe is included

in the above mentioned microarray; assigning exons to a single gene,

comprising (a) identifying exons from genomic sequence by the method

above and (b) measuring the expression of each of the exons in several

tissues and/or cell types using hybridisation to a single exon

microarrays having a probe with the exon, where a common pattern of

expression of the exons in the tissues and/or cell types indicates that

the exons should be assigned to a single gene; a peptide comprising one

of 12011 sequences, mentioned in the specification, or encoded by the

probes/open reading frames (ORF). The probes are used for gene expression

analysis, and for identifying exons in a gene, particularly using human

lung derived mRNA and for the study of lung diseases such as asthma, lung

cancer, chronic obstructive pulmonary disease (COPD), interstitial lung

disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,

tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-

Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary

histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,

Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary

dyskinesia, pulmonary hypertension and hyaline membrane disease. The

present sequence is a single exon probe open reading frame of the

invention. Note: The sequence data for this patent did not form part of

the printed specification, but was obtained in electronic format directly

from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Query Match 4.2%; Score 182; DB 6; Length 182;

Best Local Similarity 100.0%; Pred. No. 2.9e-74;

Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2545 TCCAGCTCACCTTAACCCCATGGTGGCGCTTCGGAAGATCAATGTGCTCTACTGTGC 2604

Db 182 TCCAGCTCACCTTAACCCCATGGTGGCGCTTCGGAAGATCAATGTGCTCTACTGTGC 123

QY 2605 TGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCGGTGTGAGCCAGT 2664

Db 122 TGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCGGTGTGAGCCAGT 63

QY 2665 GATGCTGATTCGTGGGAATTTATCAGAACCCAGTACTCGGGACCTGAGGCTGCAACC 2724

Db 62 GATGCTGATTCGTGGGAATTTATCAGAACCCAGTACTCGGGACCTGAGGCTGCAACC 3

QY 2725 TG 2726

Db 2 TG 1

RESULT 39

AAI14910/c

ID AAI14910 standard; DNA; 375 BP.

XX AC AAI14910;

XX DT 12-OCT-2001 (first entry)

XX DE Probe #4843 for gene expression analysis in human cervical cell sample.

XX DE Probe; human; microarray; gene expression; cervical epithelial cell;

XX DE cervical cancer; ss.

XX KW Homo sapiens.

XX OS WO200157278-A2.

XX PN 09-AUG-2001.

```
XX 30-JAN-2001; 2001WO-US000670.
PF 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 03-OCT-2000; 2000US-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488901/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human cervical epithelial cells.
XX Claim 25; SEQ ID NO 4843; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging of
XX diseases of the cervix, notably cervical cancer. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
SQ
Query Match 4.0%; Score 173; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 4.7e-70;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2592 GTCCTACTGGTGGGAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 2651
Db 375 GTCTTACTGGTGGGAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 316
QY 2652 GTCTGAGCCCAAGTATGTCGGAATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 2711
Db 315 GTCTGAGCCCAAGTATGTCGGAATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 256
QY 2712 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 2764
Db 255 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 203
RESULT 40
ABA56648/c
ID ABA56648 standard; DNA; 375 BP.
XX
XX ABA56648;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human foetal liver single exon nucleic acid probe #4953.
XX
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
XX Homo sapiens.
XX
XX WO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000669.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
```

```
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 03-OCT-2000; 2000US-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.
XX Claim 1; SEQ ID NO 4953; 639pp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
XX human gene expression in a sample derived from human foetal liver. The
XX single exon nucleic acid probes may be used for predicting, measuring and
XX displaying gene expression in samples derived from human fetal liver. The
XX present sequence is a single exon nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
SQ
Query Match 4.0%; Score 173; DB 4; Length 375;
Best Local Similarity 100.0%; Pred. No. 4.7e-70;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2592 GTCCTACTGGTGGGAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 2651
Db 375 GTCTTACTGGTGGGAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 316
QY 2652 GTCTGAGCCCAAGTATGTCGGAATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 2711
Db 315 GTCTGAGCCCAAGTATGTCGGAATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 256
QY 2712 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 2764
Db 255 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 203
RESULT 41
AAI36262/c
ID AAI36262 standard; DNA; 375 BP.
XX
XX AAI36262;
XX
XX 17-OCT-2001 (first entry)
XX
XX Probe #4948 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 03-OCT-2000; 2000US-00024263.
PR
```

PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 XX PA  
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488997/53.  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human placenta.  
 XX  
 PS Claim 25; SEQ ID NO 4948; 554pp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SENP).  
 CC The present sequence is one such probe. The probes are useful for  
 CC producing a microarray for predicting, measuring and displaying gene  
 CC expression in samples derived from human placenta. The probes are useful  
 CC for antenatal diagnosis of human genetic disorders  
 XX  
 SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;  
 Query Match 4.0%; Score 173; DB 4; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-70;  
 Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2592 GTCCTACTGGTCTGGGAAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 2651  
 DB |||||||  
 QY 375 GTCCTACTGGTCTGGGAAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 316  
 DB |||||||  
 QY 2652 GTCCTACTGGTCTGGGAAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 2711  
 DB |||||||  
 QY 315 GTCCTACTGGTCTGGGAAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 256  
 QY 2712 GAAGCTGCACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 2764  
 DB |||||||  
 DB 255 GAAGCTGCACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 203  
 RESULT 42  
 ABA46100/c  
 ID ABA46100 standard; DNA; 375 BP.  
 XX  
 AC ABA46100;  
 XX  
 DT 01-FEB-2002 (first entry)  
 XX  
 DE Human breast cell single exon nucleic acid probe #4795.  
 XX  
 KW Human; microarray; single exon probe; gene expression; breast; disease;  
 KW cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157271-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000662.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-496933/54.

PT New spatially-addressable set of single exon nucleic acid probes, useful  
 PT for measuring gene expression in sample derived from human breast,  
 PT comprises number of single exon nucleic acid probes.  
 XX  
 PS Claim 1; SEQ ID NO 4795; 327pp + Sequence Listing; English.  
 XX  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human breast and BT 474 cells. The method involves contacting the  
 CC probes with a collection of detectably labelled nucleic acids derived  
 CC from mRNA of human breast, and then measuring the label bound to each  
 CC probe of the microarray. The probes are useful for verifying the  
 CC expression of regions of genomic DNA predicted to encode proteins. They  
 CC are useful for gene discovery, and for determining predisposition and/or  
 CC prognosing breast disease. Gene expression analysis is useful for  
 CC assessing the toxicity of chemical agents on cells. The microarray of  
 CC this invention presents a far greater diversity of probes for measuring  
 CC gene expression, with far less bias than expressed sequence tag  
 CC microarrays. The method is suitable for rapid production of functional  
 CC information from genomic sequence. The present sequence is a single exon  
 CC nucleic acid probe of the invention. Note: The sequence data for this  
 CC patent did not form part of the printed specification, but was obtained  
 CC in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;  
 Query Match 4.0%; Score 173; DB 4; Length 375;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-70;  
 Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2592 GTCCTACTGGTCTGGGAAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 2651  
 DB |||||||  
 QY 375 GTCCTACTGGTCTGGGAAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 316  
 DB |||||||  
 QY 2652 GTCCTACTGGTCTGGGAAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 2711  
 DB |||||||  
 QY 315 GTCCTACTGGTCTGGGAAGATATGTCCTCCCATTTTACAATAGTTTACTACAGTGGATGCC 256  
 QY 2712 GAAGCTGCACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 2764  
 DB |||||||  
 DB 255 GAAGCTGCACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCTG 203  
 RESULT 43  
 ABA26254/c  
 ID ABA26254 standard; DNA; 375 BP.  
 XX  
 AC ABA26254;  
 XX  
 DT 23-JAN-2002 (first entry)  
 XX  
 DE Probe #4720 for gene expression analysis in human heart cell sample.  
 XX  
 KW Human; gene expression; heart; microarray; vascular system; probe;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157274-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000666.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI; 2001-48899/53.  
XX PT Single exon nucleic acid probes for analyzing gene expression in human  
XX PT hearts.  
XX PS Claim 1; SEQ ID NO 4720; 530pp; English.  
XX CC The present invention relates to single exon nucleic acid probes for  
XX CC measuring human gene expression in a sample derived from human heart. The  
XX CC present sequence is one such probe. The probes may be used for  
XX CC predicting, measuring and displaying gene expression in samples derived  
XX CC from the human heart via microarrays. By measuring gene expression, the  
XX CC probes are useful for predicting, diagnosing, grading, staging,  
XX CC monitoring and prognosing diseases of the human heart and vascular system  
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
XX CC congenital heart disease. Note: The sequence data for this patent did not  
XX CC form part of the printed specification, but was obtained in electronic  
XX CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;  
Query Match 4.0%; Score 173; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.7e-70;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2592 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 2651  
DB 375 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 316  
QY 2652 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 2711  
DB 315 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 256  
QY 2712 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 2764  
DB 255 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 203  
RESULT 44  
AAK30307/c  
ID AAK30307 standard; DNA; 375 BP.  
XX AC AAK30307;  
XX DT 06-NOV-2001 (first entry)  
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 4864.  
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;  
XX KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.  
XX OS Homo sapiens.  
XX PN WO200157276-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US000668.  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX PR 26-MAY-2000; 2000US-0207456P.  
XX PR 30-JUN-2000; 2000US-00608408.  
XX PR 03-AUG-2000; 2000US-00632366.  
XX PR 21-SEP-2000; 2000US-0234687P.  
XX PR 27-SEP-2000; 2000US-0236359P.  
XX PR 04-OCT-2000; 2000GB-00024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI; 2001-48899/53.  
XX PT Single exon nucleic acid probes for analyzing gene expression in human  
XX PT hearts.  
XX PS Claim 1; SEQ ID NO 4720; 530pp; English.  
XX CC The present invention relates to single exon nucleic acid probes for  
XX CC measuring human gene expression in a sample derived from human heart. The  
XX CC present sequence is one such probe. The probes may be used for  
XX CC predicting, measuring and displaying gene expression in samples derived  
XX CC from the human heart via microarrays. By measuring gene expression, the  
XX CC probes are useful for predicting, diagnosing, grading, staging,  
XX CC monitoring and prognosing diseases of the human heart and vascular system  
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
XX CC congenital heart disease. Note: The sequence data for this patent did not  
XX CC form part of the printed specification, but was obtained in electronic  
XX CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;  
Query Match 4.0%; Score 173; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.7e-70;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2592 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 2651  
DB 375 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 316  
QY 2652 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 2711  
DB 315 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 256  
QY 2712 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 2764  
DB 255 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 203  
RESULT 44  
AAK30307/c  
ID AAK30307 standard; DNA; 375 BP.  
XX AC AAK30307;  
XX DT 06-NOV-2001 (first entry)  
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 4864.  
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;  
XX KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.  
XX OS Homo sapiens.  
XX PN WO200157276-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US000668.  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX PR 26-MAY-2000; 2000US-0207456P.  
XX PR 30-JUN-2000; 2000US-00608408.  
XX PR 03-AUG-2000; 2000US-00632366.  
XX PR 21-SEP-2000; 2000US-0234687P.  
XX PR 27-SEP-2000; 2000US-0236359P.  
XX PR 04-OCT-2000; 2000GB-00024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI; 2001-48899/53.  
XX PT Single exon nucleic acid probes for analyzing gene expression in human  
XX PT hearts.  
XX PS Claim 1; SEQ ID NO 4720; 530pp; English.  
XX CC The present invention relates to single exon nucleic acid probes for  
XX CC measuring human gene expression in a sample derived from human heart. The  
XX CC present sequence is one such probe. The probes may be used for  
XX CC predicting, measuring and displaying gene expression in samples derived  
XX CC from the human heart via microarrays. By measuring gene expression, the  
XX CC probes are useful for predicting, diagnosing, grading, staging,  
XX CC monitoring and prognosing diseases of the human heart and vascular system  
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
XX CC congenital heart disease. Note: The sequence data for this patent did not  
XX CC form part of the printed specification, but was obtained in electronic  
XX CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;  
Query Match 4.0%; Score 173; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.7e-70;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2592 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 2651  
DB 375 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 316  
QY 2652 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 2711  
DB 315 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 256  
QY 2712 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 2764  
DB 255 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 203  
RESULT 45  
AAK04785/c  
ID AAK04785 standard; DNA; 375 BP.  
XX AC AAK04785;  
XX DT 05-NOV-2001 (first entry)  
XX DE Human brain expressed single exon probe SEQ ID NO: 4776.  
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;  
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;  
XX KW ss.  
XX OS Homo sapiens.  
XX PN WO200157275-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US000667.  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX PR 26-MAY-2000; 2000US-0207456P.  
XX PR 30-JUN-2000; 2000US-00608408.  
XX PR 03-AUG-2000; 2000US-00632366.  
XX PR 21-SEP-2000; 2000US-0234687P.  
XX PR 27-SEP-2000; 2000US-0236359P.  
XX PR 04-OCT-2000; 2000GB-00024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI; 2001-48899/53.  
XX PT Single exon nucleic acid probes for analyzing gene expression in human  
XX PT hearts.  
XX PS Claim 1; SEQ ID NO 4864; 558pp + Sequence Listing; English.  
XX CC The present invention provides a number of single exon nucleic acid  
XX CC probes which are derived from genomic sequences expressed in the human  
XX CC bone marrow. They can be used to measure gene expression in bone marrow  
XX CC samples, which may enable the improved diagnosis and treatment of cancers  
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is one of  
XX CC the probes of the invention  
XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;  
Query Match 4.0%; Score 173; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.7e-70;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2592 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 2651  
DB 375 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 316  
QY 2652 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 2711  
DB 315 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 256  
QY 2712 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 2764  
DB 255 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 203

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI; 2001-48899/53.  
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing  
XX PT gene expression in human bone marrow.  
XX PS Example 4; SEQ ID NO 4864; 558pp + Sequence Listing; English.  
XX CC The present invention provides a number of single exon nucleic acid  
XX CC probes which are derived from genomic sequences expressed in the human  
XX CC bone marrow. They can be used to measure gene expression in bone marrow  
XX CC samples, which may enable the improved diagnosis and treatment of cancers  
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is one of  
XX CC the probes of the invention  
XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;  
Query Match 4.0%; Score 173; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.7e-70;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2592 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 2651  
DB 375 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 316  
QY 2652 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 2711  
DB 315 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 256  
QY 2712 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 2764  
DB 255 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 203  
RESULT 45  
AAK04785/c  
ID AAK04785 standard; DNA; 375 BP.  
XX AC AAK04785;  
XX DT 05-NOV-2001 (first entry)  
XX DE Human brain expressed single exon probe SEQ ID NO: 4776.  
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;  
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;  
XX KW ss.  
XX OS Homo sapiens.  
XX PN WO200157275-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US000667.  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX PR 26-MAY-2000; 2000US-0207456P.  
XX PR 30-JUN-2000; 2000US-00608408.  
XX PR 03-AUG-2000; 2000US-00632366.  
XX PR 21-SEP-2000; 2000US-0234687P.  
XX PR 27-SEP-2000; 2000US-0236359P.  
XX PR 04-OCT-2000; 2000GB-00024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI; 2001-48899/53.  
XX PT Single exon nucleic acid probes for analyzing gene expression in human  
XX PT hearts.  
XX PS Example 4; SEQ ID NO 4864; 558pp + Sequence Listing; English.  
XX CC The present invention provides a number of single exon nucleic acid  
XX CC probes which are derived from genomic sequences expressed in the human  
XX CC bone marrow. They can be used to measure gene expression in bone marrow  
XX CC samples, which may enable the improved diagnosis and treatment of cancers  
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is one of  
XX CC the probes of the invention  
XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;  
Query Match 4.0%; Score 173; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.7e-70;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2592 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 2651  
DB 375 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 316  
QY 2652 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 2711  
DB 315 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 256  
QY 2712 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 2764  
DB 255 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 203  
RESULT 45  
AAK04785/c  
ID AAK04785 standard; DNA; 375 BP.  
XX AC AAK04785;  
XX DT 05-NOV-2001 (first entry)  
XX DE Human brain expressed single exon probe SEQ ID NO: 4776.  
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;  
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;  
XX KW ss.  
XX OS Homo sapiens.  
XX PN WO200157275-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US000667.  
XX PR 04-FEB-2000; 2000US-0180312P.  
XX PR 26-MAY-2000; 2000US-0207456P.  
XX PR 30-JUN-2000; 2000US-00608408.  
XX PR 03-AUG-2000; 2000US-00632366.  
XX PR 21-SEP-2000; 2000US-0234687P.  
XX PR 27-SEP-2000; 2000US-0236359P.  
XX PR 04-OCT-2000; 2000GB-00024263.  
XX PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX DR WPI; 2001-48899/53.  
XX PT Single exon nucleic acid probes for analyzing gene expression in human  
XX PT hearts.  
XX PS Example 4; SEQ ID NO 4864; 558pp + Sequence Listing; English.  
XX CC The present invention provides a number of single exon nucleic acid  
XX CC probes which are derived from genomic sequences expressed in the human  
XX CC bone marrow. They can be used to measure gene expression in bone marrow  
XX CC samples, which may enable the improved diagnosis and treatment of cancers  
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is one of  
XX CC the probes of the invention  
XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;  
Query Match 4.0%; Score 173; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.7e-70;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2592 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 2651  
DB 375 GTCTACTGCTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCC 316  
QY 2652 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 2711  
DB 315 GTCTGAGCCAGTGTGCTGGATTCGTGGAAATTTATCAGAACCCAGTACTCGGGGACT 256  
QY 2712 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 2764  
DB 255 GAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCGCTG 203

PT brains.  
XX Example 4; SEQ ID NO 4776; 650pp + Sequence Listing; English.  
PS The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is one of the probes of the  
CC invention  
XX  
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;  
Query Match 4.0%; Score 173; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.7e-70;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2592 GTCCTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 2651  
Db 375 GTCCTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 316  
QY 2652 GTCCTAGCCAGTGTGCTGGATTTCTGGAATTTATCAGAACCCAGTACTCGGGACT 2711  
Db 315 GTCCTAGCCAGTGTGCTGGATTTCTGGAATTTATCAGAACCCAGTACTCGGGACT 256  
QY 2712 GAAGGCTGCAACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTG 2764  
Db 255 GAAGGCTGCAACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTG 203  
RESULT 46  
ABS29950/c  
ID ABS29950 standard; DNA; 375 BP.  
XX  
AC ABS29950;  
XX  
DT 25-FEB-2003 (first entry)  
XX  
DE Human liver single exon probe, SEQ ID NO 4940.  
XX  
KW Human; single exon nucleic acid probe; liver; cirrhosis;  
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;  
KW coronary heart disease; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157273-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US000664.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024253.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
PI WPI; 2001-488898/53.  
XX  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human adult liver.  
XX  
PS Claim 1; SEQ ID NO 4940; 658pp; English.  
XX  
XX The invention relates to a single exon nucleic acid probe (SEN) (I) for

CC measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridises at high  
CC stringency to a nucleic acid molecule expressed in the human adult liver.  
CC (i) may be used for predicting, measuring and displaying gene expression  
CC in samples derived from human adult liver. The genes identified may be  
CC involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
CC associated with coronary heart disease. ABS25011-ABS51005 represent human  
CC liver single exon nucleic acid probes of the invention. Note: The  
CC sequence information for this patent does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;  
Query Match 4.0%; Score 173; DB 4; Length 375;  
Best Local Similarity 100.0%; Pred. No. 4.7e-70;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2592 GTCCTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 2651  
Db 375 GTCCTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCC 316  
QY 2652 GTCCTAGCCAGTGTGCTGGATTTCTGGAATTTATCAGAACCCAGTACTCGGGACT 2711  
Db 315 GTCCTAGCCAGTGTGCTGGATTTCTGGAATTTATCAGAACCCAGTACTCGGGACT 256  
QY 2712 GAAGGCTGCAACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTG 2764  
Db 255 GAAGGCTGCAACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCTG 203  
RESULT 47  
AAI04689/c  
ID AAI04689 standard; DNA; 375 BP.  
XX  
AC AAI04689;  
XX  
DT 09-OCT-2001 (first entry)  
XX  
DE Probe #4680 used to measure gene expression in human breast sample.  
XX  
KW Probe; human; breast disease; breast cancer; development disorder; ss;  
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200157270-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 29-JAN-2001; 2001WO-US000661.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
PI WPI; 2001-476286/51.  
XX  
XX Novel single exon nucleic acid probe used to measuring gene expression in  
PT a human breast.  
XX  
PS Claim 25; SEQ ID NO 4680; 322pp; English.  
XX

```
CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and non-
CC carcinoma tumours. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
Query Match 4.0%; Score 173; DB 5; Length 375;
Best Local Similarity 100.0%; Pred. No. 4.7e-70;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2592 GTCTACTGGTCTGGGAGATATGTCCTCCCAATTACAAATAGTTTACTACAGTGGATGCC 2651
Db 375 GTCTACTGGTCTGGGAGATATGTCCTCCCAATTACAAATAGTTTACTACAGTGGATGCC 316
QY 2652 GTCTGAGCCCGAGTGTGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGACT 2711
Db 315 GTCTGAGCCCGAGTGTGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGACT 256
QY 2712 GAAGGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGTCTGTGAAGGCCCTG 2764
Db 255 GAAGGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGTCTGTGAAGGCCCTG 203
RESULT 49
ABS04892/c
ID ABS04892 standard; DNA; 375 BP.
AC ABS04892;
XX
XX 19-AUG-2002 (first entry)
DE Human genome-derived single exon probe from lung SEQ ID No 4883.
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberculous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Heremansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US0000665.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX Claim 1; SEQ ID NO 4883; 634pp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of probes
XX; the novel set of probes which hybridise at high stringency to a nucleic
XX acid expressed in the human lung; measuring gene expression in a sample
XX derived from human lung, comprising (a) contacting the array with a
XX collection of detectably labeled nucleic acids derived from human lung
XX mRNA; and (b) measuring the label detectably bound to each probe of the
XX array; identifying exons in a eukaryotic genome, comprising (a)
XX algorithmically predicting at least one exon from genomic sequences of
XX the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarrays having a probe with the exon, where a common pattern of
XX expression of the exons in the tissues and/or cell types indicates that
XX the exons should be assigned to a single gene; a peptide comprising one
XX of 12011 sequences, mentioned in the specification, or encoded by the
XX probes/open reading frames (ORF). The probes are used for gene expression
XX analysis, and for identifying exons in a gene, particularly using human
XX lung derived mRNA and for the study of lung diseases such as asthma, lung
XX cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
XX disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
XX tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Heremansky-
XX Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
XX histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
XX Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
XX dyskinesia, pulmonary hypertension and hyaline membrane disease. The
XX present sequence is a single exon probe of the invention. Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
Query Match 4.0%; Score 173; DB 6; Length 375;
Best Local Similarity 100.0%; Pred. No. 4.7e-70;
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2592 GTCTACTGGTCTGGGAGATATGTCCTCCCAATTACAAATAGTTTACTACAGTGGATGCC 2651
Db 375 GTCTACTGGTCTGGGAGATATGTCCTCCCAATTACAAATAGTTTACTACAGTGGATGCC 316
QY 2652 GTCTGAGCCCGAGTGTGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGACT 2711
Db 315 GTCTGAGCCCGAGTGTGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGACT 256
QY 2712 GAAGGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGTCTGTGAAGGCCCTG 2764
Db 255 GAAGGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGTCTGTGAAGGCCCTG 203
RESULT 49
AAS69062
ID AAS69062 standard; cDNA; 2267 BP.
XX
XX AAS69062;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #4866.
```

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX Homo sapiens.  
 OS WO200175067-A2.  
 XX 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US008631.  
 XX 31-MAR-2000; 2000US-00540217.  
 XX 23-AUG-2000; 2000US-00649167.  
 XX (HYSE-) HYSEQ INC.  
 PA Drmanac RT, Liu C, Tang YT;  
 PI WPI; 2001-639362/73.  
 DR P-PSDB; ABGC4875.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 1; SEQ ID NO 4866; 103pp; English.  
 PS The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC reaction. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 2267 BP; 917 A; 455 C; 453 G; 442 T; 0 U; 0 Other;  
 SQ Query Match 3.4%; Score 148; DB 5; Length 2267;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-58;  
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3417 ATGTCAGGTAGAAATCGATACAGAGAACCTCTCAGAGGACATGTTATTACTCT 3476  
 DB 162 ATGTGAGGTAGAAATCGATACAGAGAACCTCTCAGAGGACATGTTATTACTCT 221  
 QY 3477 TCTTATTGACTATCAGTTTACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGC 3536  
 DB 222 TCTTATTGACTATCAGTTTACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGC 281  
 QY 3537 TATCAATTTTGGCTACTCTCTGACAA 3564  
 DB 282 TATCAATTTTGGCTACTCTCTGACAA 309  
 RESULT 50  
 AAI13789/c  
 ID AAI13789 standard; DNA; 402 BP.

XX AAI13789;  
 AC 12-OCT-2001 (first entry)  
 DT Probe #3722 for gene expression analysis in human cervical cell sample.  
 DE Probe; human; microarray; gene expression; cervical epithelial cell;  
 XX cervical cancer; ss.  
 KW Homo sapiens.  
 OS WO200157278-A2.  
 XX 09-AUG-2001.  
 PD 30-JAN-2001; 2001WO-US000670.  
 XX 04-FEB-2000; 2000US-0180312P.  
 XX 26-MAY-2000; 2000US-0207456P.  
 XX 30-JUN-2000; 2000US-00608408.  
 XX 03-AUG-2000; 2000US-00623266.  
 XX 21-SEP-2000; 2000US-0234687P.  
 XX 27-SEP-2000; 2000US-0236359P.  
 XX 04-OCT-2000; 2000US-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2001-488901/53.  
 DR Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human cervical epithelial cells.  
 XX Claim 25; SEQ ID NO 3722; 487pp; English.  
 PS The present invention relates to human single exon nucleic acid probes  
 CC (SENPs). The present sequence is one such probe. The SENPs are derived  
 CC from human HeLa cells. The SENPs can be used to produce a single exon  
 CC microarray, which can be used for measuring human gene expression in a  
 CC sample derived from human cervical epithelial cells. By measuring gene  
 CC expression, the probes are therefore useful in grading and/or staging of  
 CC diseases of the cervix, notably cervical cancer. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;  
 SQ Query Match 3.3%; Score 141; DB 4; Length 402;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-55;  
 Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3182 CAGCTTGCCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTG 3241  
 DB 391 CAGCTTGCCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTG 332  
 QY 3242 AGAACCTTGACCAAGCAGGACGACCTCGAGACCTGATATCTGGCTTCTACGGTATCCCA 3301  
 DB 331 AGAACCTTGACCAAGCAGGACGACCTCGAGACCTGATATCTGGCTTCTACGGTATCCCA 272  
 QY 3302 CCAATGGAGGAAATGTCAGC 3322  
 DB 271 CCAATGGAGGAAATGTCAGC 251  
 RESULT 51  
 ABA55499/c  
 ID ABA55499 standard; DNA; 402 BP.  
 XX ABA55499;  
 AC ABA55499;  
 XX





```
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-496933/54.
XX XX New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX PS Claim 1; SEQ ID NO 3722; 327pp + Sequence Listing; English.
XX CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the invention. Note: The sequence data for this
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
Query Match 3.3%; Score 141; DB 4; Length 402;
Best Local Similarity 100.0%; Pred. No. 4.3e-55;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3182 CAGCTTGCCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGTG 3241
DB 391 CAGCTTGCCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGTG 332
QY 3242 AGAACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3301
DB 331 AGAACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 332
QY 3302 CCAATGGAGGAAATGTGAGC 3322
DB 271 CCAATGGAGGAAATGTGAGC 251
RESULT 54
ABA25215/c
ID ABA25215 standard; DNA; 402 BP.
XX AC ABA25215;
XX DT 23-JAN-2002 (first entry)
XX DE Probe #3681 for gene expression analysis in human heart cell sample.
XX KW Human; gene expression; heart; microarray; vascular system; probe;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease; ss.
XX OS Homo sapiens.
XX XX
```

```
OS XX Homo sapiens.
XX PN WO200157274-A2.
XX XX
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000666.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-488899/53.
XX XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX PS Claim 1; SEQ ID NO 3681; 530pp; English.
XX CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease. Note: The sequence data for this patent did not
CC form part of the printed specification, but was obtained in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
Query Match 3.3%; Score 141; DB 4; Length 402;
Best Local Similarity 100.0%; Pred. No. 4.3e-55;
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3182 CAGCTTGCCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGTG 3241
DB 391 CAGCTTGCCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGTG 332
QY 3242 AGAACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3301
DB 331 AGAACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 332
QY 3302 CCAATGGAGGAAATGTGAGC 3322
DB 271 CCAATGGAGGAAATGTGAGC 251
RESULT 55
AAK29197/c
ID AAK29197 standard; DNA; 402 BP.
XX AC AAK29197;
XX XX
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 3754.
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX OS Homo sapiens.
XX XX
```

```
PN WO200157276-A2.
XX 09-AUG-2001.
XX 30-JAN-2001; 2001WO-US000668.
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX Example 4; SEQ ID NO 3754; 658pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX bone marrow. They can be used to measure gene expression in bone marrow
XX samples, which may enable the improved diagnosis and treatment of cancers
XX such as lymphoma, leukaemia and myeloma. The present sequence is one of
XX the probes of the invention
XX Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
XX
XX Query Match 3.3%; Score 141; DB 4; Length 402;
XX Best Local Similarity 100.0%; Pred. No. 4.3e-55;
XX Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 3182 CAGCTTGGCCATGCAACGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 3241
XX DB 391 CAGCTTGGCCATGCAACGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 332
XX QY 3242 AGAAGCTGACCAAGCAAGCACTCGAGACCTGCAATCTGCTTCTACGGTGATCCCA 3301
XX DB 331 AGAAGCTGACCAAGCAAGCACTCGAGACCTGCAATCTGCTTCTACGGTGATCCCA 272
XX QY 3302 CCAATGGAGGAAATGTCAGC 3322
XX DB 271 CCAATGGAGGAAATGTCAGC 251
XX
XX RESULT 56
XX AAK03733/c
XX ID AAK03733 standard; DNA; 402 BP.
XX AC AAK03733;
XX DT 05-NOV-2001 (first entry)
XX XX Human brain expressed single exon probe SEQ ID NO: 3724.
XX DE
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX KW ss.
XX XX Homo sapiens.
XX OS
XX XX WO200157275-A2.
XX PN
XX XX 09-AUG-2001.
XX PD
XX PF 30-JAN-2001; 2001WO-US000667.
XX XX
```

```
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX Example 4; SEQ ID NO 3724; 650pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention
XX Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
XX
XX Query Match 3.3%; Score 141; DB 4; Length 402;
XX Best Local Similarity 100.0%; Pred. No. 4.3e-55;
XX Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 3182 CAGCTTGGCCATGCAACGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 3241
XX DB 391 CAGCTTGGCCATGCAACGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 332
XX QY 3242 AGAAGCTGACCAAGCAAGCACTCGAGACCTGCAATCTGCTTCTACGGTGATCCCA 3301
XX DB 331 AGAAGCTGACCAAGCAAGCACTCGAGACCTGCAATCTGCTTCTACGGTGATCCCA 272
XX QY 3302 CCAATGGAGGAAATGTCAGC 3322
XX DB 271 CCAATGGAGGAAATGTCAGC 251
XX
XX RESULT 57
XX ABS28814/c
XX ID ABS28814 standard; DNA; 402 BP.
XX XX
XX AC ABS28814;
XX DT 25-FEB-2003 (first entry)
XX XX Human liver single exon probe, SEQ ID No 3804.
XX DE
XX DE Human; single exon nucleic acid probe; liver; cirrhosis;
XX KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX KW coronary heart disease; ss.
XX KW
XX OS Homo sapiens.
XX XX
XX XX WO200157273-A2.
XX PN
XX XX 09-AUG-2001.
XX PD
XX PF 30-JAN-2001; 2001WO-US000664.
XX XX
XX XX 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
```

PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488898/53.  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human adult liver.  
 XX Claim 1; SEQ ID NO 3804; 658pp; English.  
 XX The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridizes at high  
 CC stringency to a nucleic acid molecule expressed in the human adult liver.  
 CC (I) may be used for predicting, measuring and displaying gene expression  
 CC in samples derived from human adult liver. The genes identified may be  
 CC involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
 CC associated with coronary heart disease. ABS25011-ABS51005 represent human  
 CC liver single exon nucleic acid probes of the invention. Note: The  
 CC sequence information for this patent does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;  
 Query Match 3.3%; Score 141; DB 4; Length 402;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-55;  
 Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3182 CAGCTTGCCCAATGCAACGCGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTG 3241  
 DB 391 CAGCTTGCCCAATGCAACGCGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTG 332  
 QY 3242 AGAACCTGACCAAGCAGCAAGCACTGCGAGCATCTGCTTCTACGGTGATCCCA 3301  
 DB 331 AGAACCTGACCAAGCAGCAAGCACTGCGAGCATCTGCTTCTACGGTGATCCCA 272  
 QY 3302 CCAATGGAGGAAATGTTCAGC 3322  
 DB 271 CCAATGGAGGAAATGTTCAGC 251  
 RESULT 58  
 AA103665/c  
 ID AA103665 standard; DNA; 402 BP.  
 AC AA103665;  
 DT 09-OCT-2001 (first entry)  
 XX Probe #3656 used to measure gene expression in human breast sample.  
 XX Probe; human; breast disease; breast cancer; development disorder; ss;  
 KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
 XX Homo sapiens.  
 OS WO200157270-A2.  
 PN 09-AUG-2001.  
 XX 29-JAN-2001; 2001WO-US000661.  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207458P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632365.

PR 21-SEP-2000; 2000US-0234587P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-476286/51.  
 XX Novel single exon nucleic acid probe used to measuring gene expression in  
 PT a human breast.  
 XX Claim 25; SEQ ID NO 3656; 322pp; English.  
 XX The present invention relates to novel single exon nucleic acid probes.  
 CC The present sequence is one such probe. The probes are useful for  
 CC measuring human gene expression in a human breast sample, where the probe  
 CC hybridizes at high stringency to a nucleic acid expressed in the human  
 CC breast. The probes are useful for predicting, diagnosing, grading,  
 CC staging, monitoring and prognosing diseases of the human breast,  
 CC particularly those diseases with polygenic aetiology. The diseases  
 CC include: breast cancer, disorders of development, inflammatory diseases  
 CC of the breast, fibrocystic changes, proliferative breast disease and non-  
 CC carcinoma tumours. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;  
 Query Match 3.3%; Score 141; DB 5; Length 402;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-55;  
 Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3182 CAGCTTGCCCAATGCAACGCGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTG 3241  
 DB 391 CAGCTTGCCCAATGCAACGCGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTG 332  
 QY 3242 AGAACCTGACCAAGCAGCAAGCACTGCGAGCATCTGCTTCTACGGTGATCCCA 3301  
 DB 331 AGAACCTGACCAAGCAGCAAGCACTGCGAGCATCTGCTTCTACGGTGATCCCA 272  
 QY 3302 CCAATGGAGGAAATGTTCAGC 3322  
 DB 271 CCAATGGAGGAAATGTTCAGC 251  
 RESULT 59  
 ABS03749/c  
 ID ABS03749 standard; DNA; 402 BP.  
 XX ABS03749;  
 AC ABS03749;  
 DT 19-AUG-2002 (first entry)  
 XX Human genome-derived single exon probe from lung SEQ ID No 3740.  
 XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberculous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW hyaline membrane disease.  
 XX Homo sapiens.  
 OS WO200186003-A2.  
 PN 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US000665.  
XX 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2002-114183/15.  
XX Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples.  
XX Claim 1; SEQ ID NO 3740; 634pp; English.  
XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of probes  
CC; the novel set of probes which hybridize at high stringency to a nucleic  
CC acid expressed in the human lung; measuring gene expression in a sample  
CC derived from human lung, comprising (a) contacting the array with a  
CC collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of the  
CC array; identifying exons in a eukaryotic genome, comprising (a)  
CC algorithmically predicting at least one exon from genomic sequences of  
CC the eukaryote; and (b) detecting specific hybridization of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridization to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene expression  
CC analysis, and for identifying exons in a gene, particularly using human  
CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
CC dyskinesia, pulmonary hypertension and bialine membrane disease. The  
CC present sequence is a single exon probe of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences  
SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;  
Query Match 3.3%; Score 141; DB 6; Length 402;  
Best Local Similarity 100.0%; Pred. No. 4.3e-55;  
Matches 141; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3182 CAGCTTGCCAAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTG 3241  
DB 391 CAGCTTGCCAAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTG 332  
QY 3242 AGAAGCTGACACAGCAAGCAAGCACTGGAGACCTGATATCTGGCTTCTACGGTGATCCCA 3301  
DB 331 AGAAGCTGACACAGCAAGCAAGCACTGGAGACCTGATATCTGGCTTCTACGGTGATCCCA 272

QY 3302 CCAATGGAGGGAATGTCAGC 3322  
DB 271 CCAATGGAGGGAATGTCAGC 251  
RESULT 60  
AAI23009/c  
ID AAI23009 standard; DNA; 137 BP.  
XX AC AAI23009;  
XX 12-OCT-2001 (first entry)  
XX Probe #12942 for gene expression analysis in human cervical cell sample.  
DE Probe; human; microarray; gene expression; cervical epithelial cell;  
XX KW cervical cancer; ss.  
XX OS Homo sapiens.  
XX PN WO200157278-A2.  
XX PD 09-AUG-2001.  
XX PF 30-JAN-2001; 2001WO-US000670.  
XX 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234687P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GB-00024263.  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
XX WPI; 2001-488901/53.  
XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human cervical epithelial cells.  
XX Claim 25; SEQ ID NO 12942; 487pp; English.  
XX The present invention relates to human single exon nucleic acid probes  
CC (SENPs). The present sequence is one such probe. The SENPs are derived  
CC from human HeLa cells. The SENPs can be used to produce a single exon  
CC microarray, which can be used for measuring human gene expression in a  
CC sample derived from human cervical epithelial cells. By measuring gene  
CC expression, the probes are therefore useful in grading and/or staging of  
CC diseases of the cervix, notably cervical cancer. Note: The sequence data  
CC for this patent did not form part of the printed specification, but was  
CC obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;  
SQ Query Match 3.2%; Score 137; DB 4; Length 137;  
Best Local Similarity 100.0%; Pred. No. 3.2e-53;  
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3185 CTTGCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTG 3244  
DB 137 CTTGCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAGTG 78  
QY 3245 ACCTGACACACAGGCAAGCACTGGAGACCTGATATCTGGCTTCTACGGTGATCCCA 3304  
DB 77 ACCTGACACACAGGCAAGCACTGGAGACCTGATATCTGGCTTCTACGGTGATCCCA 18  
QY 3305 ATGAGGGAATGTCAG 3321

```

Db      17 ATGGAGGGAATGTCTAG 1
RESULT 61
ABR68099/c
ID   ABA68099 standard; DNA; 137 BP.
XX
XX
AC   ABA68099;
XX
DT   01-FEB-2002 (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #16404.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000669.
XX
PR   04-FEB-2000; 2000US-0180312P.
XX
PR   26-MAY-2000; 2000US-0207456P.
XX
PR   30-JUN-2000; 2000US-00608408.
XX
PR   03-AUG-2000; 2000US-00632366.
XX
PR   21-SEP-2000; 2000US-0234687P.
XX
PR   27-SEP-2000; 2000US-0236359P.
XX
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human foetal liver.
XX
PS   Claim 4; SEQ ID NO 16404; 639pp + Sequence Listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for measuring
CC   human gene expression in a sample derived from human foetal liver. The
CC   single exon nucleic acid probes may be used for predicting, measuring and
CC   displaying gene expression in samples derived from human foetal liver. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Query Match      3.2%; Score 137; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.2e-53;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   3185 CTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244
Db      137 CTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78

QY   3245 ACCTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCA 3304
Db      77 ACCTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCA 18

QY   3305 ATGGAGGGAATGTCTAG 3321
Db      17 ATGGAGGGAATGTCTAG 1

RESULT 62
AAI48309/c
ID   AAI48309 standard; DNA; 137 BP.
XX
XX
AC   AAI48309;
XX
DT   17-OCT-2001 (first entry)
XX
DE   Probe #16995 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
XX   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US000663.
XX
PR   04-FEB-2000; 2000US-0180312P.
XX
PR   26-MAY-2000; 2000US-0207456P.
XX
PR   30-JUN-2000; 2000US-00608408.
XX
PR   03-AUG-2000; 2000US-00632366.
XX
PR   21-SEP-2000; 2000US-0234687P.
XX
PR   27-SEP-2000; 2000US-0236359P.
XX
PR   04-OCT-2000; 2000GB-00024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for analyzing
PT   gene expression in human placenta.
XX
PS   Claim 25; SEQ ID NO 16995; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENPs).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders
XX
SQ   Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Query Match      3.2%; Score 137; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.2e-53;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY   3185 CTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244
Db      137 CTTGCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78

QY   3245 ACCTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCA 3304
Db      77 ACCTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTATCCACCA 18

QY   3305 ATGGAGGGAATGTCTAG 3321
Db      17 ATGGAGGGAATGTCTAG 1

RESULT 63
ABA50165/c
ID   ABA50165 standard; DNA; 137 BP.
XX
XX
AC   ABA50165;
XX
DT   01-FEB-2002 (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #8860.
XX
KW   Human; microarray; single exon probe; gene expression; breast; disease;

```

```

KW cancer; ss.
XX
XX Homo sapiens.
OS
XX WO200157271-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000662.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes, useful
XX PT for measuring gene expression in sample derived from human breast,
XX PT comprises number of single exon nucleic acid probes.
XX
XX Claim 4; SEQ ID NO 8860; 327pp + Sequence Listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and BT 474 cells. The method involves contacting the
XX CC probes with a collection of detectably labelled nucleic acids derived
XX CC from mRNA of human breast, and then measuring the label bound to each
XX CC probe of the microarray. The probes are useful for verifying the
XX CC expression of regions of genomic DNA predicted to encode proteins. They
XX CC are useful for gene discovery, and for determining predisposition and/or
XX CC assessing the toxicity of chemical agents on cells. The microarray of
XX CC this invention presents a far greater diversity of probes for measuring
XX CC gene expression, with far less bias than expressed sequence tag
XX CC microarrays. The method is suitable for rapid production of functional
XX CC information from genomic sequence. The present sequence is a single exon
XX CC nucleic acid probe of the invention. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;
Query Match 3.2%; Score 137; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.2e-53;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3185 CTTGCCAATGCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244
DB 137 CTTGCCAATGCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78
QY 3245 ACCTGACCAAGGCAAGCACTGGAGACCTGCAATCTGGCTTCTACGGTATCCCA 3304
DB 77 ACCTGACCAAGGCAAGCACTGGAGACCTGCAATCTGGCTTCTACGGTATCCCA 18
QY 3305 ATGGAGGGAATGTCTAG 3321
DB 17 ATGGAGGGAATGTCTAG 1
RESULT 64
ABA35129/C
ID ABA35129 standard; DNA; 137 BP.
XX
XX ABA35129;
AC
```

```

XX
XX 23-JAN-2002 (first entry)
XX
XX Probe #13595 for gene expression analysis in human heart cell sample.
XX
XX Human; gene expression; heart; microarray; vascular system; probe;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease; ss.
XX
XX Homo sapiens.
XX OS
XX WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000666.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX
XX 26-MAY-2000; 2000US-0207456P.
XX
XX 30-JUN-2000; 2000US-00608408.
XX
XX 03-AUG-2000; 2000US-00632366.
XX
XX 21-SEP-2000; 2000US-0234687P.
XX
XX 27-SEP-2000; 2000US-0236359P.
XX
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-498899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts.
XX
XX Claim 4; SEQ ID NO 13595; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart. The
XX CC present sequence is one such probe. The probes may be used for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from the human heart via microarrays. By measuring gene expression, the
XX CC probes are useful for predicting, diagnosing, grading, staging,
XX CC monitoring and prognosing diseases of the human heart and vascular system
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX CC congenital heart disease. Note: The sequence data for this patent did not
XX CC form part of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;
Query Match 3.2%; Score 137; DB 4; Length 137;
Best Local Similarity 100.0%; Pred. No. 3.2e-53;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3185 CTTGCCAATGCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244
DB 137 CTTGCCAATGCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78
QY 3245 ACCTGACCAAGGCAAGCACTGGAGACCTGCAATCTGGCTTCTACGGTATCCCA 3304
DB 77 ACCTGACCAAGGCAAGCACTGGAGACCTGCAATCTGGCTTCTACGGTATCCCA 18
QY 3305 ATGGAGGGAATGTCTAG 3321
DB 17 ATGGAGGGAATGTCTAG 1
RESULT 65
AAK42239/C
ID AAK42239 standard; DNA; 137 BP.
XX
XX AAK42239;
AC
```

DT XX 06-NOV-2001 (first entry)  
 DE XX Human bone marrow expressed single exon probe SEQ ID NO: 16796.  
 DE XX Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157276-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 XX 30-JAN-2001; 2001WO-US000668.  
 XX  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-488900/53.  
 DR  
 XX Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human bone marrow.  
 XX  
 XX Example 4; SEQ ID NO 16796; 658pp + Sequence Listing; English.  
 PS  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of  
 CC the probes of the invention  
 XX  
 XX Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;  
 SQ  
 Query Match 3.2%; Score 137; DB 4; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-53;  
 Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3185 CTGCGCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244  
 Db 137 CTGCGCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78  
 QY 3245 ACCTGACCACAGGCAAGCACTGCGAGACTGCGAGACTGCGATATCTGGCTTCTACGGTATCCACCA 3304  
 Db 77 ACCTGACCACAGGCAAGCACTGCGAGACTGCGATATCTGGCTTCTACGGTATCCACCA 18  
 QY 3305 ATGGAGGGAATGTCTAG 3321  
 Db 17 ATGGAGGGAATGTCTAG 1  
 RESULT 66  
 AAK16486/c  
 ID AAK16486 standard; DNA; 137 BP.  
 XX  
 AC AAK16486;  
 XX  
 XX 05-NOV-2001 (first entry)  
 DT  
 DE Human brain expressed single exon probe SEQ ID NO: 16477.  
 XX  
 XX Human; brain expressed exon; gene expression analysis; probe; microarray;  
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;  
 KW ss.

XX Homo sapiens.  
 OS  
 XX WO200157275-A2.  
 PN  
 XX 09-AUG-2001.  
 PD  
 XX 30-JAN-2001; 2001WO-US000667.  
 XX  
 XX 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2001-483446/52.  
 DR  
 XX Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains.  
 PT  
 XX Example 4; SEQ ID NO 16477; 650pp + Sequence Listing; English.  
 PS  
 XX The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system,  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is one of the probes of the  
 CC invention  
 XX  
 XX Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;  
 SQ  
 Query Match 3.2%; Score 137; DB 4; Length 137;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-53;  
 Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3185 CTGCGCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244  
 Db 137 CTGCGCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78  
 QY 3245 ACCTGACCACAGGCAAGCACTGCGAGACTGCGAGACTGCGATATCTGGCTTCTACGGTATCCACCA 3304  
 Db 77 ACCTGACCACAGGCAAGCACTGCGAGACTGCGATATCTGGCTTCTACGGTATCCACCA 18  
 QY 3305 ATGGAGGGAATGTCTAG 3321  
 Db 17 ATGGAGGGAATGTCTAG 1  
 RESULT 67  
 ABS41845/c  
 ID ABS41845 standard; DNA; 137 BP.  
 XX  
 AC ABS41845;  
 XX  
 XX 25-FEB-2003 (first entry)  
 DT  
 DE Human liver single exon probe, SEQ ID No 16835.  
 XX  
 XX Human; single exon nucleic acid probe; liver; cirrhosis;  
 KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;  
 KW coronary heart disease; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157273-A2.  
 XX



09-AUG-2001.  
30-JAN-2001; 2001WO-US000664.  
04-FEB-2000; 2000US-0180312P.  
26-MAY-2000; 2000US-0207456P.  
30-JUN-2000; 2000US-00608408.  
03-AUG-2000; 2000US-00632366.  
21-SEP-2000; 2000US-0234687P.  
27-SEP-2000; 2000US-0236359P.  
04-OCT-2000; 2000GB-00024263.  
(MOLE-) MOLECULAR DYNAMICS INC.  
Penn SG, Hanzel DK, Chen W, Rank DR;  
WPI; 2001-488898/53.  
Human genome-derived single exon nucleic acid probes useful for analyzing  
gene expression in human adult liver.  
Claim 4; SEQ ID NO 16935; 656pp; English.  
The invention relates to a single exon nucleic acid probe (SEN) (I) for  
measuring human gene expression in a sample derived from human adult  
liver, comprising one of 13109 defined nucleotide sequences given in the  
specification (or complements/ fragments). The probe hybridises at high  
stringency to a nucleic acid molecule expressed in the human adult liver.  
(I) may be used for predicting, measuring and displaying gene expression  
in samples derived from human adult liver. The genes identified may be  
involved in genetic liver diseases such as cirrhosis,  
hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
associated with coronary heart disease. ABS25011-ABS51005 represent human  
liver single exon nucleic acid probes of the invention. Note: The  
sequence information for this patent does not appear in the printed  
specification but was obtained in electronic format directly from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences  
Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;  
Query Match 3.28; Score 137; DB 4; Length 137;  
Best Local Similarity 100.0%; Pred. No. 3.2e-53;  
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3185 CTTGCCAATGCAACGGCCACAGTAATGTCATCAATCAGACGATCTGTGAGAAGTGTGAGA 3244  
Db 137 CTTGCCAATGCAACGGCCACAGTAATGTCATCAATCAGACGATCTGTGAGAAGTGTGAGA 78  
QY 3245 ACCTGACCACAGGCAAGCATTGGAGACCTGCGATATCTGGCTTCTACGGTATCCACCA 3304  
Db 77 ACCTGACCACAGGCAAGCATTGGAGACCTGCGATATCTGGCTTCTACGGTATCCACCA 18  
QY 3305 ATGGAGGGAAATGTCAG 3321  
Db 17 ATGGAGGGAAATGTCAG 1  
RESULT 68  
AAI08671/c  
ID AAI08671 standard; DNA; 137 BP.  
XX AAI08671;  
XX AC  
XX XX  
XX DT (first entry)  
DE 09-OCT-2001  
XX Probe #8662 used to measure gene expression in human breast sample.  
XX Probe; human; breast disease; breast cancer; development disorder; ss;  
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX Homo sapiens.  
XX OS  
XX OS  
XX WO200157270-A2  
XX NN

KW Primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease; open reading frame; ORF.

OS Homo sapiens.

PN W0200186003-A2.

XX 15-NOV-2001.

PF 30-JAN-2001; 2001WO-US000665.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2002-114183/15.

DR Spatially-addressable set of single exon nucleic acid probes, used to

PT measure gene expression in human lung samples.

XX Claim 4; SEQ ID NO 16280; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of probes  
 CC; the novel set of probes which hybridise at high stringency to a nucleic  
 CC acid expressed in the human lung; measuring gene expression in a sample  
 CC derived from human lung, comprising (a) contacting the array with a  
 CC collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of the  
 CC array; identifying exons in a eukaryotic genome, comprising (a)  
 CC algorithmically predicting at least one exon from genomic sequences of  
 CC the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene expression  
 CC analysis, and for identifying exons in a gene, particularly using human  
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
 CC Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
 CC present sequence is a single exon probe open reading frame of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

XX Query Match 3.2%; Score 137; DB 6; Length 137;

XX Best Local Similarity 100.0%; Pred. No. 3.2e-53;

Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3185 CTTGCCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGA 3244  
 Db 137 CTTGCCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGA 78  
 QY 3245 ACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 3304  
 Db 77 ACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 18  
 QY 3305 ATGGAGGGAATGTCTAG 3321  
 Db 17 ATGGAGGGAATGTCTAG 1

RESULT 70

AAAL23548/c

ID AAL23548 standard; cDNA; 482 BP.

XX AAL23548;

XX 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 16005.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US000798.

XX 14-JAN-2000; 2000US-0176077P.

XX 14-MAR-2000; 2000US-0189167P.

XX 24-MAR-2000; 2000US-0192099P.

XX 29-MAR-2000; 2000US-0193480P.

XX 15-MAY-2000; 2000US-0205230P.

XX 09-JUN-2000; 2000US-0211315P.

XX 25-JUL-2000; 2000US-0220534P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer.

XX Claim 1; Page 2922; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides  
 CC (AAU07544-AAU26789) and methods of assessing whether a patient is  
 CC afflicted with breast cancer by examining the correlation between the  
 CC expression of certain markers and the cancerous state of breast cells.  
 CC The polynucleotides and encoded polypeptides are potential markers for  
 CC detecting, diagnosing, monitoring, characterising treating and  
 CC potentially preventing breast cancer. The polynucleotides and encoded  
 CC polypeptides are also useful for isolating compounds with cytostatic  
 CC activity

XX Sequence 482 BP; 106 A; 133 C; 126 G; 117 T; 0 U; 0 Other;

XX Query Match 3.0%; Score 130; DB 4; Length 482;

XX Best Local Similarity 100.0%; Pred. No. 6e-50;

XX Matches 130; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 2848 ATGTGGTGCAGCAACATGAAGCAGTGTGTGACCTCCAAATGCCTATGTGGCCTCTTCCCT 2907

XX 130 ATGTGGTGCAGCAACATGAAGCAGTGTGTGACCTCCAAATGCCTATGTGGCCTCTTCCCT 71

[illegible]

```

RESULT 71
AAS77314
ID AAS77314 standard; cDNA; 503 BP.
XX
XX AAS77314;
XX
XX 13-FEB-2002 (first entry)
XX
XX DNA encoding novel human diagnostic protein #13118.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
XX OS Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX PR P-PSDB; ABG13127.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PR responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX
XX Claim 1; SEQ ID NO 13118; 103pp; English.

```

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences, AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

Query Match 1.4%; Score 62; DB 5; Length 503;  
Best Local Similarity 100.0%; Pred. No. 3.8e-18;

	Matches	62;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	3802	ATTGCCTTCTCTCAGACACAGCAATTTATGGACCTGGTACAGTCTTCGTGACTTTCTTC								3861
Db	130	ATTGCCTTCTCTCAGACACAGCAATTTATGGACCTGGTACAGTCTTCGTGACTTTCTTC								189
Qy	3862	AG 3863								
Db	190	AG 191								

RESULT 72	
ADBS3249	
ID	ADB53249 standard; DNA; 8739 BP.
XX	
AC	ADB53249;
XX	
XX	
04-DEC-2003	(first entry)
DT	
XX	
DE	Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3791.
XX	
XX	toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW	toxicity marker; toxicity progression; drug screening;
KW	primary rat hepatocyte toxicity modelling; gene; ds.
XX	
OS	Rattus norvegicus.
XX	
XX	
FN	WO2003065993-A2.
XX	
PD	14-AUG-2003.
XX	
XX	
PF	04-FEB-2003; 2003WO-US003482.
XX	
XX	
PR	04-FEB-2002; 2002US-0353171P.
PR	13-MAR-2002; 2002US-0363344P.
PR	08-APR-2002; 2002US-0370248P.
PR	10-APR-2002; 2002US-0371134P.
PR	10-APR-2002; 2002US-0371135P.
PR	10-APR-2002; 2002US-0371150P.
PR	11-APR-2002; 2002US-0371413P.
PR	19-APR-2002; 2002US-0373601P.
PR	19-APR-2002; 2002US-0373602P.
PR	22-APR-2002; 2002US-0374139P.
PR	08-MAY-2002; 2002US-0378370P.
PR	09-MAY-2002; 2002US-0378652P.
PR	09-MAY-2002; 2002US-0378653P.
PR	09-MAY-2002; 2002US-0378665P.
PR	09-JUL-2002; 2002US-0394230P.
PR	09-JUL-2002; 2002US-0394253P.
PR	04-SEP-2002; 2002US-0407688P.
PR	28-JAN-2003; 2003US-0442900P.
XX	
PA	(GENE-) GENE LOGIC INC.
XX	

XX Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;  
PI Elashoff N;  
XX WPI; 2003-731472/69.  
DR

AA Determining if a compound induces a toxic effect on a tissue or cell, for  
 PT identifying hepatotoxic compounds, comprises comparing a gene expression  
 PT profile of a tissue or cell sample to a database of Tox mean and non-Tox  
 PT mean values.

XX  
PS Claim 44: SEQ ID NO 3791; 874pp; English.

XX The present invention describes a method for determining whether a  
CC compound induces a toxic effect on a tissue or cell. The method comprises  
CC preparing a gene expression profile of a tissue or cell sample exposed to  
CC the compound, and comparing the gene expression profile to a database  
CC comprising data or information on the Tox mean and non-Tox mean value.  
CC The method is useful for predicting or identifying at least one toxic  
CC effect, particularly hepatotoxicity, of a test or unknown compound. The

CC genes listed in the specification are useful as diagnostic or toxicity  
 CC markers for the prediction or identification of the physiological state  
 CC of tissue or cell sample that has been exposed to a compound, or to  
 CC identify or predict the toxic effects of a compound or an agent. These  
 CC may also be used as markers for monitoring toxicity progression or for  
 CC drug screening. The present sequence represents a primary rat hepatocyte  
 CC toxicity modelling related gene sequence from the present invention.

XX  
 XX Sequence 8739 BP; 2288 A; 2115 C; 2109 G; 2227 T; 0 U; 0 Other;

Query Match 1.4%; Score 62; DB 9; Length 8739;  
 Best Local Similarity 100.0%; Pred. No. 3.7e-18;  
 Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 GACTGGTGGGAGCAATGCCAGCACTCGCGGGGCGCGTTCAGACTAAGTGGATCTTCT 426  
 |||||  
 Db 436 GCGTGGTGGGAGCAATGCCAGCACTCGCGGGGCGCGTTCAGACTAAGTGGATCTTCT 495  
 |||||

QY 427 GG 428  
 ||  
 Db 496 GG 497

RESULT 74  
 AABN42673  
 ID AABN42673 standard; DNA; 60 BP.

XX AC AABN42673;  
 XX DT 15-JUL-2002 (first entry)

XX DE Human spliced transcript detection oligonucleotide SEQ ID NO:15421.  
 XX KW Human; mouse; rat; splice transcript; detection; RNA transcript;  
 XX KW splice variant; transcriptome; oligonucleotide library; ss.  
 XX OS Homo sapiens.  
 XX PN WC200210449-A2.  
 XX PD 07-FEB-2002.  
 XX PF 20-JUL-2001; 2001WO-IB001903.  
 XX PR 28-JUL-2000; 2000US-0221607P.  
 XX PR 02-MAY-2001; 2001US-0287724P.  
 XX PA (COMP-) COMPUGEN INC.  
 XX PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
 XX WPI; 2002-257383/30.

XX New oligonucleotide libraries comprising oligonucleotides which  
 PT selectively hybridize to mRNAs transcribed from a transcription unit of a  
 PT genome, useful for detecting tissue-, pathology-, and developmental-  
 PT specific genes.

XX Example 1; SEQ ID NO 15421; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting  
 CC messenger RNAs that populate a (sub-)transcriptome, where the (sub-  
 CC )transcriptome comprises messenger RNAs transcribed from multiple  
 CC transcription units that populate a genome. The library comprises several  
 CC oligonucleotides, each capable of hybridising selectively to a set of  
 CC messenger RNAs transcribed from a given transcription unit of the genome,  
 CC which encodes one or more messenger RNA splice variants. The  
 CC oligonucleotide libraries are useful for detecting mRNAs from a  
 CC biological sample, in expression profiling studies, in qualitatively or  
 CC quantitatively characterising the corresponding transcriptome, and in  
 CC detecting RNA transcripts and splice variants of human or animal  
 CC transcriptomes. The libraries may also be used as specialised mini  
 CC libraries to detect transcripts of a sub-transcriptome under a particular

CC biological or pathological state, and so allowing the detection of tissue  
 CC - and pathology-specific genes such as those genes only expressed in  
 CC specific tissue under a specific pathological condition; to detect  
 CC developmental specific genes; and to detect RNA transcripts and splice  
 CC variants of a transcriptome of a patient suffering from a particular  
 CC disorder. AABN42673 to AABN42679 represent oligonucleotide sequences from  
 CC rats, humans and mice, which are used in the exemplification of the  
 CC present invention. N.B. The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX  
 XX Sequence 60 BP; 19 A; 14 C; 13 G; 14 T; 0 U; 0 Other;

Query Match 1.4%; Score 60; DB 6; Length 60;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-17;  
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3547 GTGGCTACTCTCGACGAACAACAGGATTTGGACATGTTTCATCAATGCTCCAGAAT 3606  
 |||||  
 Db 1 GTGGCTACTCTCGACGAACAACAGGATTTGGACATGTTTCATCAATGCTCCAGAAT 60  
 |||||

RESULT 74  
 AABN42673  
 ID AABN42673 standard; DNA; 90050 BP.

XX AC AABN42673;  
 XX DT 08-JUN-2000 (first entry)

XX DE Wild type (C57BL/6J) murine mahogany protein genomic sequence c110/111.  
 XX KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;  
 XX KW weight regulation; cell therapy; body weight disorder; cachexia;  
 XX KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
 XX KW hyperphagia; Anticobesity; antianorexic; anticachexic; ds.  
 XX OS Mus sp.  
 XX PN WC2000005373-A2.  
 XX PD 03-FEB-2000.  
 XX PF 21-JUL-1999; 99WO-US016484.  
 XX PR 21-JUL-1998; 98US-0093630P.  
 XX PR 20-OCT-1998; 98US-0104978P.  
 XX PR 05-FEB-1999; 99US-00245041.  
 XX PA (MILL-) MILLENIUM PHARM INC.  
 XX PI Moore K, Nagle DL;  
 XX WPI; 2000-195103/17.

XX New human and murine mahogany genes, useful, e.g. for diagnosis and  
 PT treatment of body weight disorders.

XX Example; Fig 3d; 188pp; English.

XX This sequence represents a murine mahogany gene of the invention. The  
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iii) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in (cell-  
 CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (II) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (II) for diagnosis, prognosis or monitoring of treatment; to  
 CC evaluate (II)-expressing cells intended for cell therapy; and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the

CC mg polypeptide) are used to identify agents (A) that modulate mg  
CC activity. (A) are potentially useful for the treatment of body weight  
CC disorders, particularly obesity, cachexia or anorexia, or other  
CC conditions associated with the mg gene such as hyperpigmentation,  
CC hyperphagia and disorders that result in increased metabolic rate  
XX  
SQ Sequence 90050 BP; 25307 A; 18193 C; 19454 G; 27096 T; 0 U; 0 Other;  
Query Match 1.3%; Score 55; DB 3; Length 90050;  
Best Local Similarity 100.0%; Pred. No. 6.9e-15;  
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3130 TCCAGCATGCTCTAGAGGACGAGCAGATACACTGCTCTTCATTCTCACTGTCACG 3184  
DB 31840 TCCAGCATGCTCTAGAGGACGAGCAGATACAACTGGTCTTTCATTCTCACTGTCACG 31894

RESULT 75  
AAZ91918  
ID AAZ91918 standard; cDNA; 1051 BP.  
XX  
AC AAZ91918;  
XX  
DT 08-JUN-2000 (first entry)  
XX  
DE Murine mahogany protein coding sequence akml004.  
XX  
KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;  
KW weight regulation; cell therapy; body weight disorder; cachexia;  
KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.  
XX  
OS Mus sp.  
XX  
FN WO200005373-A2.  
XX  
PD 03-FEB-2000.  
XX  
PF 21-JUL-1999; 95WO-US01:6484.  
XX  
PR 21-JUL-1998; 98US-0093630P.  
PR 20-OCT-1998; 98US-0104978P.  
PR 05-FEB-1999; 99US-00245041.  
XX  
PA (MILL-) MILLENIUM PHARM INC.  
XX  
PI Moore X, Nagle DL;  
XX  
DR WPI; 2000-195103/17.  
DR P-PSDB; AAY81805.  
XX  
PT New human and murine mahogany genes, useful, e.g. for diagnosis and  
PT treatment of body weight disorders.  
XX  
PS Claim 1; Fig 9a; 188pp; English.  
XX  
CC This sequence represents a murine mahogany gene of the invention. The  
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming  
CC therapeutics; (iii) as a source of diagnostic probes and primers for  
CC detecting expression of mg genes or mutations, regulatory defects, in  
CC this gene, or for isolation of related sequences; and (iv) in (cell-  
CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to  
CC identify other (extra)cellular products involved in weight regulation,  
CC and to screen for agents that disrupt interaction between (ii) and other  
CC macromolecules. The Ab are used to detect abnormal levels (or function)  
CC of (ii) (for diagnosis, prognosis or monitoring of treatment); to  
CC evaluate (iii)-expressing cells intended for cell therapy, and as  
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
CC mg polypeptide) are used to identify agents (A) that modulate mg  
CC activity. (A) are potentially useful for the treatment of body weight  
CC disorders, particularly obesity, cachexia or anorexia, or other  
CC conditions associated with the mg gene such as hyperpigmentation,

CC hyperphagia and disorders that result in increased metabolic rate  
XX  
SQ Sequence 1051 BP; 306 A; 243 C; 313 G; 189 T; 0 U; 0 Other;  
Query Match 1.2%; Score 53; DB 3; Length 1051;  
Best Local Similarity 100.0%; Pred. No. 6.1e-14;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 367 GGCTGGTGGGCGAGCAATGCCAGCACTCGGGGGCGCTTCAGACTAACTGG 419  
DB 571 GGCTGGTGGGCGAGCAATGCCAGCACTCGGGGGCGCTTCAGACTAACTGG 623

Search completed: March 1, 2004, 13:39:57  
Job time : 1480 secs

***This Page Blank (uspto)***



c 89 22 0.5 459 9 US-09-864-761-2182 Sequence 2182, Ap  
90 22 0.5 459 9 US-09-981-876-36 Sequence 36, Appl  
91 22 0.5 459 10 US-09-148-545-36 Sequence 36, Appl  
92 22 0.5 480 10 US-09-770-961-950 Sequence 950, App  
93 22 0.5 554 9 US-09-864-761-7357 Sequence 7357, Ap  
c 94 22 0.5 585 14 US-10-029-386-4261 Sequence 4261, Ap  
c 95 22 0.5 607 12 US-10-425-114-22348 Sequence 22348, A  
c 96 22 0.5 616 14 US-10-029-386-22934 Sequence 22934, A  
c 97 22 0.5 616 15 US-10-242-535A-1932 Sequence 1932, Ap  
c 98 22 0.5 654 14 US-10-156-761-2741 Sequence 2741, Ap  
c 99 22 0.5 665 12 US-10-424-599-97537 Sequence 97537, A  
c 100 22 0.5 775 12 US-10-424-599-16675 Sequence 16675, A

ALIGNMENTS

RESULT 1  
US-09-893-238-14  
; Sequence 14, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893, 238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245, 041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093, 630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104, 978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 8589  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-893-238-14

Query Match 88.0%; Score 3774; DB 9; Length 8589;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3974; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 313 TGTGTCAACGGCGTGCCTGCAACCCCTGGCACCAGTGCCTGCTGCCCGCGCGGTGG 372  
DB 310 TGTGTCAACGGCGTGCCTGCAACCCCTGGCACCAGTGCCTGCTGCCCGCGCGGTGG 369  
QY 373 GTGGCGAGCAATGCCAGCACTGGCGGGCGCTTCAGACTAATGATCTTCTGGGTTT 432  
DB 370 GTGGCGAGCAATGCCAGCACTGGCGGGCGCTTCAGACTAATGATCTTCTGGGTTT 429  
QY 433 GTGACAGATGACCTGGGAAATTAATAAACAAGTGCAGTGCCTGCTCATTAAGGA 492  
DB 430 GTGACAGATGACCTGGGAAATTAATAAACAAGTGCAGTGCCTGCTCATTAAGGA 489  
QY 493 CAGCCAAATAGAAATAGCACTTCGTTTCAATCAATTTGCTACAGAGTGTAGTGGAC 552  
DB 490 CAGCCAAATAGAAATAGCACTTCGTTTCAATCAATTTGCTACAGAGTGTAGTGGAC 549  
QY 553 CATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTGCAATTTAGTGGC 612  
DB 550 CATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTGCAATTTAGTGGC 609  
QY 613 CTCAATGTTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTGTTGCCACATCAGGT 672  
DB 610 CTCAATGTTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTGTTGCCACATCAGGT 669  
QY 673 TATGCTTGTGCTGCAATTTTATGATGCTGCTTATATTTGACTGGATTAATTAATTA 732

DB 670 TATGCTTGTGCTGCAATTTTATGATGCTGCTTATAATTTGACTGGAATTAATTAATTA 729  
QY 733 TACAGTTTTGATATGTTGCCAATAAATTAAGTCTCAGCCGAGGAGAGTGAAGATCAGTAAT 792  
DB 730 TACAGTTTTGATATGTTGCCAATAAATTAAGTCTCAGCCGAGGAGAGTGAAGATCAGTAAT 789  
QY 793 AGCAGCGAAACTGTTCAATGTGAATGTTTCTGAAAACTGGAAAGGTGAAGCATGTGACATT 852  
DB 790 AGCAGCGAATGTTGTAATGTGAATGTTTCTGAAAACTGGAAAGGTGAAGCATGTGACATT 849  
QY 853 CCTCACTGTACAGACAACCTGCTGCTTCTCTATCGAGCATCTGCAATTTCAAGTCAATGTC 912  
DB 850 CCTCACTGTACAGACAACCTGCTGCTTCTCTATCGAGCATCTGCAATTTCAAGTCAATGTC 909  
QY 913 AGAGATGCTCTCTGCTTCTCAGACTGGCAGGCTCTCGATGTTTCACTTCTCTGACAGCT 972  
DB 910 AGAGATGCTCTCTGCTTCTCAGACTGGCAGGCTCTCGATGTTTCACTTCTCTGACAGCT 969  
QY 973 AACAGTCAATTTGGACTCGAGAGAAATTTCTAATTTAAAGTCTCCAGAGCATCTCAT 1032  
DB 970 AACAGTCAATTTGGACTCGAGAGAAATTTCTAATTTAAAGTCTCCAGAGCATCTCAT 1029  
QY 1033 AAAGCTGTGCTCAATGGAACAATTTATGCGGTGTTTGGAGGATATATGTTCAACCACTCA 1092  
DB 1030 AAAGCTGTGCTCAATGGAACAATTTATGCGGTGTTTGGAGGATATATGTTCAACCACTCA 1089  
QY 1093 GATTATAACATGTTTCTAGCGTATGACCTTCTTAGGAGTGGCTTCCACTAAACCGT 1152  
DB 1090 GATTATAACATGTTTCTAGCGTATGACCTTCTTAGGAGTGGCTTCCACTAAACCGT 1149  
QY 1153 TCTGTGAACAATGCTGTTTGTAGATATGCTGCTTCTTGGCATTATACAAGGATAAAAAT 1212  
DB 1150 TCTGTGAACAATGCTGTTTGTAGATATGCTGCTTCTTGGCATTATACAAGGATAAAAAT 1209  
QY 1213 TACATGATGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAGTGTGAGATTTT 1272  
DB 1210 TACATGATGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAGTGTGAGATTTT 1269  
QY 1273 CACATTCATTAATGAGTCACTGGGTGTTGTGACCCCTAAGGCAAGGAGCAGTATCAGTG 1332  
DB 1270 CACATTCATTAATGAGTCACTGGGTGTTGTGACCCCTAAGGCAAGGAGCAGTATCAGTG 1329  
QY 1333 GTTGGCACTCTGCACACATTTCTTACACTGAAGAATGCGCAGTGTGCTGCTGCTATC 1392  
DB 1330 GTTGGCACTCTGCACACATTTCTTACACTGAAGAATGCGCAGTGTGCTGCTGCTATC 1389  
QY 1393 TTTGTCATGCTGCTCTCTATGGAATATATAGCAATGTGCAGGAATATGATTTGGATAG 1452  
DB 1390 TTTGTCATGCTGCTCTCTATGGAATATATAGCAATGTGCAGGAATATGATTTGGATAG 1449  
QY 1453 AACACATGAGTATATTAACACACCCAGGCTGCTTGTGCAAGGGGGTTACGGCCATAGC 1512  
DB 1450 AACACATGAGTATATTAACACACCCAGGCTGCTTGTGCAAGGGGGTTACGGCCATAGC 1509  
QY 1513 AGTGTTTACGACCATAGGACACAGGCGCTTATAGTTCATGTTGGCTACAAGCTTTTCACT 1572  
DB 1510 AGTGTTTACGACCATAGGACACAGGCGCTTATAGTTCATGTTGGCTACAAGCTTTTCACT 1569  
QY 1573 GCATTAATGATCGGCTTGCAGATGATCTTACCGATATGATGTGATACCCAGATGTGG 1632  
DB 1570 GCATTAATGATCGGCTTGCAGATGATCTTACCGATATGATGTGATACCCAGATGTGG 1629  
QY 1633 ACCATTTCTTAAGGACAGCGGATTTTCCGTTTCTTGTGCAACAGCTGTGATAGTGGGA 1692  
DB 1630 ACCATTTCTTAAGGACAGCGGATTTTCCGTTTCTTGTGCAACAGCTGTGATAGTGGGA 1689  
QY 1693 ACCATGCTGTTTGGGGGAAACACACATGACATCTCTATGAGCCATGGCCGCAAA 1752  
DB 1690 ACCATGCTGTTTGGGGGAAACACACATGACATCTCTATGAGCCATGGCCGCAAA 1749  
QY 1753 TGCTTCTTTCAGATTTTATGCGCTTATGACATTTGCTGACCGCTGTGCTGCTGCTTCCC 1812  
DB 1750 TGCTTCTTTCAGATTTTATGCGCTTATGACATTTGCTGACCGCTGTGCTGCTGCTTCCC 1809



1813 AGACCTGATCTCCACCATGATGTCAAAGATTTGGCCATTTCAGCAGTCTTTACACACAGC 1872  
1810 AGACCTGATCTCCACCATGATGTCAAAGATTTGGCCATTTCAGCAGTCTTTACACACAGC 1869  
1873 ACCATGATGTTGGTTCGATGTTTCAATAGTCTCTCTCAGCAGCATCTCTGGTATTCAAC 1932  
1870 ACCATGATGTTGGTTCGATGTTTCAATAGTCTCTCTCAGCAGCATCTCTGGTATTCAAC 1929  
1933 TCGGAAACAGTGTGATGGCATCGAGTGAAGCCGCTGTTTAGCAGAGACCTGGTATT 1992  
1930 TCGGAAACAGTGTGATGGCATCGAGTGAAGCCGCTGTTTAGCAGAGACCTGGTATT 1989  
1993 CGGTGTGTGGAAACACAGAGGTCTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAA 2052  
1990 CGGTGTGTGGAAACACAGAGGTCTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAA 2049  
2053 CAAGAAGAAAGTTAAATCAGAAATGTTTTTCNAAAGAACTCTTGACCAATGACAGATGT 2112  
2050 CAAGAAGAAAGTTAAATCAGAAATGTTTTTCNAAAGAACTCTTGACCAATGACAGATGT 2109  
2113 GACCAGCACACAGATGTTTACAGCTGTACAGCCAAACCAACCACTGSCCACTGTCGAAT 2172  
2110 GACCAGCACACAGATGTTTACAGCTGTACAGCCAAACCAACCACTGSCCACTGTCGAAT 2169  
2173 GACCATGTGTCTCCAGGAACCAAGCTGTCTGAGAGCCAGATCTCCATTTTTAGGTAT 2232  
2170 GACCATGTGTCTCCAGGAACCAAGCTGTCTGAGAGCCAGATCTCCATTTTTAGGTAT 2229  
2233 GAGAAATGCCCCAAGGATACCCCATGTACTACTGTAAACCAAGAGACCACTGCGAGGC 2292  
2230 GAGAAATGCCCCAAGGATACCCCATGTACTACTGTAAACCAAGAGACCACTGCGAGGC 2289  
2293 TGTGCTTGGACCAAGAACTGCACTGGGAGCCCGGAATCAGAGTGCATTTGCCCTGCC 2352  
2290 TGTGCTTGGACCAAGAACTGCACTGGGAGCCCGGAATCAGAGTGCATTTGCCCTGCC 2349  
2353 GAAATATCTGTGTCATGCTGTCATTTGGTGGAAACTCTATGTTGAAATTTACTACT 2412  
2350 GAAATATCTGTGTCATGCTGTCATTTGGTGGAAACTCTATGTTGAAATTTACTACT 2409  
2413 GCCAAGAGAAATATGCAATGCTTAAATGTTCTGTAGGAACCAACATGCCCCTTTGGCT 2472  
2410 GCCAAGAGAAATATGCAATGCTTAAATGTTCTGTAGGAACCAACATGCCCCTTTGGCT 2469  
2473 TCTCTTCAACCCAGAGAAAGGTAGAAATTTGTCCTTAAAGCAGCTGCGAATAATGCACTCA 2532  
2470 TCTCTTCAACCCAGAGAAAGGTAGAAATTTGTCCTTAAAGCAGCTGCGAATAATGCACTCA 2529  
2533 TCTCAGAGCATGTCAGAGCTCACTTAAACCCATGGTGGGCTTCGAGAGATCAATGTG 2592  
2530 TCTCAGAGCATGTCAGAGCTCACTTAAACCCATGGTGGGCTTCGAGAGATCAATGTG 2589  
2593 TCTTACTGGTGTGGGAGATATGTCCTCCATTTTACAAATAGTTTACTACAGTGGATGCCG 2652  
2590 TCTTACTGGTGTGGGAGATATGTCCTCCATTTTACAAATAGTTTACTACAGTGGATGCCG 2649  
2653 TCTGAGCCAGTGTGATGATTTCTGTGGAATTTTATCAGAAACCCAGTACTCGGGAGCTG 2712  
2650 TCTGAGCCAGTGTGATGATTTCTGTGGAATTTTATCAGAAACCCAGTACTCGGGAGCTG 2709  
2713 AAGGCTGCAACTGCATCAACCCACTCAATGGTAGTGTGTGAAAGGCCCTGCAAAACCAAC 2772  
2710 AAGGCTGCAACTGCATCAACCCACTCAATGGTAGTGTGTGAAAGGCCCTGCAAAACCAAC 2769  
2773 AGTCTAAGCAGTGTCCGGAACCAATGTCCTTGAAGACAGCATGTGGAGATTCACCAAGC 2832  
2770 AGTCTAAGCAGTGTCCGGAACCAATGTCCTTGAAGACAGCATGTGGAGATTCACCAAGC 2829  
2833 GGCAGCTCTGAGTGTGTCAGCAACATGAGCAGTGTGTGAGCTCCCATGCTTAT 2892  
2830 GGCAGCTCTGAGTGTGTCAGCAACATGAGCAGTGTGTGAGCTCCCATGCTTAT 2889

2893 GTGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGATATACGATGAGCACCTGCCCCCT 2952  
2890 GTGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGATATACGATGAGCACCTGCCCCCT 2949  
2953 GAAATTTGTTAGGCTTACTGTACCTGTAGTCTATTGCTTTGGAGCAACAGGCTGTGGCTGG 3012  
2950 GAAATTTGTTAGGCTTACTGTACCTGTAGTCTATTGCTTTGGAGCAACAGGCTGTGGCTGG 3009  
3013 TGTACTGTATCCAGCAATACTGCGAAAGGGAATGCATAGAGGTTCTTATAAGAGCA 3072  
3010 TGTACTGTATCCAGCAATACTGCGAAAGGGAATGCATAGAGGTTCTTATAAGAGCA 3069  
3073 GTCAAGATGCTTTCGCAAGCCCTTACAGAAATTTCTATCCACAGCCCTGTCTCAATTC 3132  
3070 GTCAAGATGCTTTCGCAAGCCCTTACAGAAATTTCTATCCACAGCCCTGTCTCAATTC 3129  
3133 AGCATGTGTCTAGAGGACAGCAGATACAACCTGGTCTTTCTATTCATCTGTCCAGTTGCCAA 3192  
3130 AGCATGTGTCTAGAGGACAGCAGATACAACCTGGTCTTTCTATTCATCTGTCCAGTTGCCAA 3189  
3193 TCGAAGCCGACAGTAAATGCAATCAATCAGAGCATCTGTGAGAACTGTGAGAACTGAC 3252  
3190 TCGAAGCCGACAGTAAATGCAATCAATCAGAGCATCTGTGAGAACTGTGAGAACTGAC 3249  
3253 ACAGGCAAGCATCTGCGAGACCTGCGATATCTGGCTTCTACGGTGATCCCAATGGAGG 3312  
3250 ACAGGCAAGCATCTGCGAGACCTGCGATATCTGGCTTCTACGGTGATCCCAATGGAGG 3309  
3313 AATGTGCGCATGCAAGTGCATGCGCAAGTGCATGCGTCTTACGGTGATCCCAATGGAGG 3372  
3310 AATGTGCGCATGCAAGTGCATGCGCAAGTGCATGCGTCTTACGGTGATCCCAATGGAGG 3369  
3373 TGTCTCTGACCAACCAAGGGCGTCAAGGGGACGAGTGCAGCTATGTGAGTAGAAAT 3432  
3370 TGTCTCTGACCAACCAAGGGCGTCAAGGGGACGAGTGCAGCTATGTGAGTAGAAAT 3429  
3433 CGATACCAAGGAAACCTCTCAGAGGAAATGTTTATATCTCTTCTTATTGACTATCAG 3492  
3430 CGATACCAAGGAAACCTCTCAGAGGAAATGTTTATATCTCTTCTTATTGACTATCAG 3489  
3493 TTTACCTTTAGTCTATCCAGGAAATGATCGCTATTTACAGCTATCAATTTTGGCT 3552  
3490 TTTACCTTTAGTCTATCCAGGAAATGATCGCTATTTACAGCTATCAATTTTGGCT 3549  
3553 ACTCTGACGAAACCAACAGGGATTTGGACATGTTTCAATCAATGCTCCAAAGAAATTCAC 3612  
3550 ACTCTGACGAAACCAACAGGGATTTGGACATGTTTCAATCAATGCTCCAAAGAAATTCAC 3609  
3613 CTCACATCACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCT 3672  
3610 CTCACATCACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCT 3669  
3673 GTTGTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTTAATGAGAAATTTGAT 3732  
3670 GTTGTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTTAATGAGAAATTTGAT 3729  
3733 TTTTCGAAACCCCAATATCACTTTCTTTGTTTATGTCAGTAATTTACCTGGCCCATC 3792  
3730 TTTTCGAAACCCCAATATCACTTTCTTTGTTTATGTCAGTAATTTACCTGGCCCATC 3789  
3793 AAAATTCAGATTTGCTTCTCAGACACCAATTTTATGGAACCTGGTACAGTTCTCGTG 3852  
3790 AAAATTCAGATTTGCTTCTCAGACACCAATTTTATGGAACCTGGTACAGTTCTCGTG 3849  
3853 ACTTCTTTCAGTGTGCTTCTCTTTGCTCTCTGTTGGCTGTCTGTGGTGGAGATCAAA 3912  
3850 ACTTCTTTCAGTGTGCTTCTCTTTGCTCTCTTGTGCTGTCTGTGGTGGAGATCAAA 3909  
3913 CAAAGTGTGTTGGCTCTCAGAGCTAGAGAGCACTCTTCTGAGAGATGCAACAGATGCC 3972  
3910 CAAAGTGTGTTGGCTCTCAGAGCTAGAGAGCACTCTTCTGAGAGATGCAACAGATGCC 3969  
3973 AGCCGTCCCTTTGCTCTGTAAATGTCGCTTTGGAACACAGATGAGGAGCTCTCTGATCTT 4032

Db	3970	AGCGTCCCTTGGCTCTGTAAATGTGCGCTTGGAAACAGATGAGGAGCCTCTGATCTT	4029
Qy	4033	ATTGGGGGAGTATAAGAGCTGTTCCCAACCCATTGCACTGGAGCGGTGTTTGGCAAC	4092
Db	4030	ATTGGGGGAGTATAAGAGCTGTTCCCAACCCATTGCACTGGAGCGGTGTTTGGCAAC	4089
Qy	4093	AAAGCCGCTGTCCTCTCTGCTGTTGTGAGGCTCCCTCGAGGCCCTGGGTGGCATCCCTCCT	4152
Db	4090	AAAGCCGCTGTCCTCTCTGCTGTTGTGAGGCTCCCTCGAGGCCCTGGGTGGCATCCCTCCT	4149
Qy	4153	CCTGGGAGTCAAGTCTGTTGGGCGAGCGCCCTGGTGGACATTTCTGAGCAGATGCCG	4212
Db	4150	CCTGGGAGTCAAGTCTGTTGGGCGAGCGCCCTGGTGGACATTTCTGAGCAGATGCCG	4209
Qy	4213	ATAGTGTACAAGGAGAAGTCAAGGAGCGGTGAGAAACCGGAGCAGAGCGCCCTGCACAG	4272
Db	4210	ATAGTGTACAAGGAGAAGTCAAGGAGCGGTGAGAAACCGGAGCAGAGCGCCCTGCACAG	4269
Qy	4273	CCTGGGACCTGCATCTGA	4290
Db	4270	CCTGGGACCTGCATCTGA	4287

RESULT 2

US-09-893-238-16

; Sequence 16, Application US/09893238

; Patent No. US20020150973A1

; GENERAL INFORMATION:

; APPLICANT: Moore, K.

; APPLICANT: Nagle, D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND

; FILE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY

; FILE REFERENCE: 7853-237

; CURRENT APPLICATION NUMBER: US/09/893,238

; CURRENT FILING DATE: 2001-06-27

; PRIOR APPLICATION NUMBER: 09/245,041

; PRIOR FILING DATE: 1999-02-05

; PRIOR APPLICATION NUMBER: 60/093,630

; PRIOR FILING DATE: 1998-07-21

; PRIOR APPLICATION NUMBER: 60/104,978

; PRIOR FILING DATE: 1998-10-20

; NUMBER OF SEQ ID NOS: 129

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 16

; LENGTH: 4072

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-893-238-16

Query Match	76.6%	Score 3285;	DB 9;	Length 4072;
Best Local Similarity	99.9%	Pred. No. 0;	Mismatches	4;
Matches 3485;	Conservative	0;	Indels	0;
Gaps	0;			

Qy	313	TGTGTCAACGGCGGTGCGTCAACCCCTGGCAACCGGCGAGTGGCTGCTCCCGCGCGGTGG	372
Db	310	TGTGTCAACGGCGGTGCGTCAACCCCTGGCAACCGGCGAGTGGCTGCTCCCGCGCGGTGG	369
Qy	373	GTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTCACTGGATCTTCTGGGTTT	432
Db	370	GTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTCACTGGATCTTCTGGGTTT	429
Qy	433	GTGACAGATGACCTGGAAATATAAATACAAACGAAGTGCACGTGGCTCATTTGAAGGA	492
Db	430	GTGACAGATGACCTGGAAATATAAATACAAACGAAGTGCACGTGGCTCATTTGAAGGA	489
Qy	493	CAGCCAAATAGAAATAGACTTCGTTTCAATCTTTCATAGAGTGTAGTGGGAC	552
Db	490	CAGCCAAATAGAAATAGACTTCGTTTCAATCTTTCATAGAGTGTAGTGGGAC	549
Qy	553	CATTATATGTTTATGATGGGAGTCAATTTATGACCGGTAGTGTGCTGATTTAGTGGC	612
Db	550	CATTATATGTTTATGATGGGAGTCAATTTATGACCGGTAGTGTGCTGATTTAGTGGC	609

Qy	613	CTCATTGTTCTTGAGAGATGGCAATGAGACTGTCTCCTGAGGTTGTTGCCACATCAGGT	672
Db	610	CTCATTGTTCTTGAGAGATGGCAATGAGACTGTCTCCTGAGGTTGTTGCCACATCAGGT	669
Qy	673	TATGCTTCTGCTGCTATTTTATGATGCTGCTGTTAATTTGACTGGATTTAATTTACT	732
Db	670	TATGCTTCTGCTGCTATTTTATGATGCTGCTGTTAATTTGACTGGATTTAATTTACT	729
Qy	733	TACAGTTTGTATGATGCTGCTCAAAATACTGCTCAGGCCGAGGAGAGTGAAGATCAGTAAT	792
Db	730	TACAGTTTGTATGCTGCTCAAAATACTGCTCAGGCCGAGGAGAGTGAAGATCAGTAAT	789
Qy	793	AGCAGCGAACTGTTGAATGTAATGTTCTGAAACTGGAAGAGTGAAGATGAGCAAT	852
Db	790	AGCAGCGATCTGTTGAATGTAATGTTCTGAAACTGGAAGAGTGAAGATGAGCAAT	849
Qy	853	CCTCACTGTACAGACAACTGTGTTTCTCATCGAGGCATCTGCAATTCAGTGAATGTC	912
Db	850	CCTCACTGTACAGACAACTGTGTTTCTCATCGAGGCATCTGCAATTCAGTGAATGTC	909
Qy	913	AGAGGATGCTCCTGCTTCTCAGACTGGCAGGCTCCTGATGTTTCTGTTTCTGTTACCACT	972
Db	910	AGAGGATGCTCCTGCTTCTCAGACTGGCAGGCTCCTGATGTTTCTGTTTCTGTTACCACT	969
Qy	973	AACAGTCAATTTTGGACTCGAGAGAAATTTCTAACTTTAAAGCTCCCGAGAGCATCTCAT	1032
Db	970	AACAGTCAATTTTGGACTCGAGAGAAATTTCTAACTTTAAAGCTCCCGAGAGCATCTCAT	1029
Qy	1033	AAAGCTGTGTTCAATGGAACAACTATGTTGGTGTGTTGGAGGATATATGTTCAACCACTCA	1092
Db	1030	AAAGCTGTGTTCAATGGAACAACTATGTTGGTGTGTTGGAGGATATATGTTCAACCACTCA	1089
Qy	1093	GATTATAACAGTGTCTAGCGTATGACCTTCTCTAGGAGTGGCTTCCACTAAACCGT	1152
Db	1090	GATTATAACAGTGTCTAGCGTATGACCTTCTCTAGGAGTGGCTTCCACTAAACCGT	1149
Qy	1153	TCTGTGAACAATGTGTTGTTAGATATGTTGATGTTCTTTGGCATTTATCAAGCAATAAAT	1212
Db	1150	TCTGTGAACAATGTGTTGTTAGATATGTTGATGTTCTTTGGCATTTATCAAGCAATAAAT	1209
Qy	1213	TACATGTATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGATGAGTGGATTTT	1272
Db	1210	TACATGTATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGATGAGTGGATTTT	1269
Qy	1273	CACATTCATATGAGTCAATGGTGTGTTGACCTTAAAGGCAAGGAGCAGTATGCAATG	1332
Db	1270	CACATTCATATGAGTCAATGGTGTGTTGACCTTAAAGGCAAGGAGCAGTATGCAATG	1329
Qy	1333	GTGGGCACTCTGCACACATTTTACACTGAAGAAATGGCCGAGTGGTTCATGCTGGTCAATC	1392
Db	1330	GTGGGCACTCTGCACACATTTTACACTGAAGAAATGGCCGAGTGGTTCATGCTGGTCAATC	1389
Qy	1393	TTTGTGCTACCTGCTCTCTATGATATATAGCAATGTGAGGAAATATGATTTGGATTAAG	1452
Db	1390	TTTGTGCTACCTGCTCTCTATGATATATAGCAATGTGAGGAAATATGATTTGGATTAAG	1449
Qy	1453	AACACATGGAGTATATTACACACCCAGGCTGCTTGTGCAAGGGGGTTACGGCCATAGC	1512
Db	1450	AACACATGGAGTATATTACACACCCAGGCTGCTTGTGCAAGGGGGTTACGGCCATAGC	1509
Qy	1513	AGTGTGTTACGACATAGAACCCAGGCGCTTATACGTTTCAATGTTGGCTACAAGCTTTTCA	1572
Db	1510	AGTGTGTTACGACATAGAACCCAGGCGCTTATACGTTTCAATGTTGGCTACAAGCTTTTCA	1569
Qy	1573	GCCAAATAGTACCGCTTGCAGATGATCTCTACCGATATGATGTTGGATACCCAGATGTTG	1632
Db	1570	GCCAAATAGTACCGCTTGCAGATGATCTCTACCGATATGATGTTGGATACCCAGATGTTG	1629
Qy	1633	ACCATTCATTAAGGACAGCCGATTTTCCGTTTCTGTTGACACAGCTGTGATGAGTGGGA	1692
Db	1630	ACCATTCATTAAGGACAGCCGATTTTCCGTTTCTGTTGACACAGCTGTGATGAGTGGGA	1689

QY 1693 ACCATGCTGGTGTGGGGAACACACACAATGACACATCTATGAGCCATGGGCCAAA 1752  
DB 1690 ACCATGCTGGTGTGGGGAACACACACAATGACACATCTATGAGCCATGGGCCAAA 1749  
QY 1753 TGCTTCTCTTCAGATTTTCATGSCCTATGACATTTCCCTGTGACCGCTGGTCACTGCTTCCC 1812  
DB 1750 TGCTTCTCTTCAGATTTTCATGSCCTATGACATTTCCCTGTGACCGCTGGTCACTGCTTCCC 1809  
QY 1813 AGACTGATCTCCACCATGATGTCACACAGATTTGGCCATTCAGCAGTCCTTACACACAGC 1872  
DB 1810 AGACTGATCTCCACCATGATGTCACACAGATTTGGCCATTCAGCAGTCCTTACACACAGC 1869  
QY 1873 ACCATGATGTTGTCGGTGTGTTTCAATAGTCTCCTCCTCAGCAGATCTCCTGATGAA 1932  
DB 1870 ACCATGATGTTGTCGGTGTGTTTCAATAGTCTCCTCCTCAGCAGATCTCCTGATGAA 1929  
QY 1933 TCGAACAAGTGTGATGCGCATTCGAGTGAAGCCCTGTTTTCAGCAGGACCTGATTT 1992  
DB 1930 TCGAACAAGTGTGATGCGCATTCGAGTGAAGCCCTGTTTTCAGCAGGACCTGATTT 1989  
QY 1993 CGGTGTGTGGAACACAGAGGTCTGCTCAGTGTATCTCGTGGGGCTGGCAACTGATGAA 2052  
DB 1990 CGGTGTGTGGAACACAGAGGTCTGCTCAGTGTATCTCGTGGGGCTGGCAACTGATGAA 2049  
QY 2053 CAAGAAGAAAGTTAAATCAGAAATGTTTTCAAAAGAACTCTGTGACCATGACAGATGT 2112  
DB 2050 CAAGAAGAAAGTTAAATCAGAAATGTTTTCAAAAGAACTCTGTGACCATGACAGATGT 2109  
QY 2113 GACCAGCACACAGATTTGTACAGCTGTACAGCCACACCAATGATCTGCCACTGGTGAAT 2172  
DB 2110 GACCAGCACACAGATTTGTACAGCTGTACAGCCACACCAATGATCTGCCACTGGTGAAT 2169  
QY 2173 GACCAATTTGTGCCAGGAACACAGCTGTCTCAGAAGCCAGATCTCCATTTTAGGTAT 2232  
DB 2170 GACCAATTTGTGCCAGGAACACAGCTGTCTCAGAAGCCAGATCTCCATTTTAGGTAT 2229  
QY 2233 GAGAAATTTGCCAGGATTAACCCATGTAATCTGTAACAGAGACACAGCTGAGGAGC 2292  
DB 2230 GAGAAATTTGCCAGGATTAACCCATGTAATCTGTAACAGAGACACAGCTGAGGAGC 2289  
QY 2293 TGTGCTGTGACAGAACTGCGAGTGGAGCCCGGAATCAGGAGTCAATTTGCCCTGCC 2352  
DB 2290 TGTGCTGTGACAGAACTGCGAGTGGAGCCCGGAATCAGGAGTCAATTTGCCCTGCC 2349  
QY 2353 GAAATATCTGTGCAATTTGGTGTGCAATTTGGTGTGAAATCTATGTTTGAATTTACTACT 2412  
DB 2350 GAAATATCTGTGCAATTTGGTGTGCAATTTGGTGTGAAATCTATGTTTGAATTTACTACT 2409  
QY 2413 GCCAAGGAAATTAATGCAATGTAATTTGTTGTAGGACACCAATGCCCTTTGGCT 2472  
DB 2410 GCCAAGGAAATTAATGCAATGTAATTTGTTGTAGGACACCAATGCCCTTTGGCT 2469  
QY 2473 TCTCTTACACCCAGAAAGGTAGAATTTGCTTAAAGCAGCTGCGAATAATGCAATGCA 2532  
DB 2470 TCTCTTACACCCAGAAAGGTAGAATTTGCTTAAAGCAGCTGCGAATAATGCAATGCA 2529  
QY 2533 TCTCAGACATGTCAGACCTCACTTAACCCATGAGTGGCTTCCGAAATCAATGTC 2592  
DB 2530 TCTCAGACATGTCAGACCTCACTTAACCCATGAGTGGCTTCCGAAATCAATGTC 2589  
QY 2593 TCCTACTGGTCTGGGAAGATGTCCTCCATTTTACAATAGTTTACTACAGTGGATGCCG 2652  
DB 2590 TCCTACTGGTCTGGGAAGATGTCCTCCATTTTACAATAGTTTACTACAGTGGATGCCG 2649  
QY 2653 TCTGAGCCAGTGTGAGATTTCTGTGGAATTTTATCAGAACCCAGTACTCTGCGGAGCTG 2712  
DB 2650 TCTGAGCCAGTGTGAGATTTCTGTGGAATTTTATCAGAACCCAGTACTCTGCGGAGCTG 2709  
QY 2713 AAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGTGAAAGCCCTGCAAAACAC 2772  
DB 2710 AAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGTGAAAGCCCTGCAAAACAC 2769  
QY 2773 AGTCTAAGCAGTGCAGACACCACTGTCCTTGTAGGACAGCATGTGGAGATTCACACAGC 2832

DB 2770 AGTCTAAGCAGTGCAGACACCACTGTCCTTGGAGACAGCATGTGGAGTTGCACAGC 2829  
QY 2833 GGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAAGCAGTGTGTGACTTCCAAATGCCAT 2892  
DB 2830 GGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAAGCAGTGTGTGACTTCCAAATGCCAT 2889  
QY 2893 GTGGCCCTCTCCCTTTTGGCCAGTGTGAAATGATATACGATGAGACCTGCCCCCT 2952  
DB 2890 GTGGCCCTCTCTCCCTTTTGGCCAGTGTGAAATGATATACGATGAGACCTGCCCCCT 2949  
QY 2953 GAAATTTGTTTTCAGGCTTACTGTACTTGTAGTCAATCTTGGAGCAACCAAGCTGTGGCTG 3012  
DB 2950 GAAATTTGTTTTCAGGCTTACTGTACTTGTAGTCAATCTTGGAGCAACCAAGCTGTGGCTG 3009  
QY 3013 TGTACTGATCCAGCAATACTGCGCAAGGAAATGCATAGAGGTTCTCTATAAAGACCA 3072  
DB 3010 TGTACTGATCCAGCAATACTGCGCAAGGAAATGCATAGAGGTTCTCTATAAAGACCA 3069  
QY 3073 GTGAAGATGCTTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATTTCC 3132  
DB 3070 GTGAAGATGCTTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATTTCC 3129  
QY 3133 AGCATGTGTAGAGACAGCAGATACAACTGCTCTTTCATTTCACTGTCCAGCTTGCCTAA 3192  
DB 3130 AGCATGTGTAGAGACAGCAGATACAACTGCTCTTTCATTTCACTGTCCAGCTTGCCTAA 3189  
QY 3193 TGCACCGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAACTGTGAGAACCTGACC 3252  
DB 3190 TGCACCGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAACTGTGAGAACCTGACC 3249  
QY 3253 ACAGGCAAGCAGTGCAGACCTTGCATATCTGCTCTTCTACGCTGATCCCAATGGAGG 3312  
DB 3250 ACAGGCAAGCAGTGCAGACCTTGCATATCTGCTCTTCTACGCTGATCCCAATGGAGG 3309  
QY 3313 AAATGTGACCCATGCAAGTGCATGCGCACGCTCTCTGTGCAACACCAACACGGGCAAG 3372  
DB 3310 AAATGTGACCCATGCAAGTGCATGCGCACGCTCTCTGTGCAACACCAACACGGGCAAG 3369  
QY 3373 TGCTTCTGACCAACCAAGGGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAT 3432  
DB 3370 TGCTTCTGACCAACCAAGGGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAT 3429  
QY 3433 CGATACCAAGGAAACCTCTCTCAGAGAACATGTTATATCTCTCTTATTTGACTATCAG 3492  
DB 3430 CGATACCAAGGAAACCTCTCTCAGAGAACATGTTATATCTCTCTTATTTGACTATCAG 3489  
QY 3493 TTCACCTTTTGTCTATCCAGGAGATGATCGCTATTTACAGACTATCAATTTTGGCT 3552  
DB 3490 TTCACCTTTTGTCTATCCAGGAGATGATCGCTATTTACAGACTATCAATTTTGGCT 3549  
QY 3553 ACTCTGACGAAACAAACAGGGATTTGGACATGTTTCATCAATGCTCTCCAAAGATTTCAAC 3612  
DB 3550 ACTCTGACGAAACAAACAGGGATTTGGACATGTTTCATCAATGCTCTCCAAAGATTTCAAC 3609  
QY 3613 CTCACATCACTGGCTGCGCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGATGCT 3672  
DB 3610 CTCACATCACTGGCTGCGCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGATGCT 3669  
QY 3673 GTTGTTCAAAACCAACATTAAGGAGTCAAAAGATAGTTTCTCTAATGAGAAATTTGAT 3732  
DB 3670 GTTGTTCAAAACCAACATTAAGGAGTCAAAAGATAGTTTCTCTAATGAGAAATTTGAT 3729  
QY 3733 TTTGCAACACCCCAATATCACTTCTTGTGTTTATGTAGTAAATTTCACTGGGCCATC 3792  
DB 3730 TTTGCAACACCCCAATATCACTTCTTGTGTTTATGTAGTAAATTTCACTGGGCCATC 3789  
QY 3793 AAAATTCAG 3801  
DB 3790 AAAATTCAG 3798

RESULT 3

US-09-893-238-18  
; Sequence 18, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-05-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 2625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-893-238-18

Query Match 42.8%; Score 1837; DB 9; Length 2625;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2037; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	313	TGTGTCACAGCGGTGCTGCAACCCCTGGCAGCGGCGAGTGCCTGCCCGCGGCTGG	372
DB	310	TGTGTCACAGCGGTGCTGCAACCCCTGGCAGCGGCGAGTGCCTGCCCGCGGCTGG	369
QY	373	GTGGCGAGCAATCCAGCACTCGGGGGCGGCTTCAGACTAACTGATCTTTGGGTTT	432
DB	370	GTGGCGAGCAATCCAGCACTCGGGGGCGGCTTCAGACTAACTGATCTTTGGGTTT	429
QY	433	GTGACAGTGCAGCTGGGAATTAATAACAAAGAGTGCAGTGGCTTCATTGAGGA	492
DB	430	GTGACAGTGCAGCTGGGAATTAATAACAAAGAGTGCAGTGGCTTCATTGAGGA	489
QY	493	CAGCAAAATAGAAATACAGACTTCGTTTCAATCAATTTTGCTACAGAGTGTAGTTGGAC	552
DB	490	CAGCAAAATAGAAATACAGACTTCGTTTCAATCAATTTTGCTACAGAGTGTAGTTGGAC	549
QY	553	CATTATATGTTATGATGGGACTCAATTTATGCAACGGCTAGTTGCTGATTTAGTGGC	612
DB	550	CATTATATGTTATGATGGGACTCAATTTATGCAACGGCTAGTTGCTGATTTAGTGGC	609
QY	613	CTCATTGTTCTGAGAGATGGCAATGAGACTGTCCTGAGGTTGTTGCCACATCAGGT	672
DB	610	CTCATTGTTCTGAGAGATGGCAATGAGACTGTCCTGAGGTTGTTGCCACATCAGGT	669
QY	673	TATGCTTGCTGCAATTTTTTTAGTGTGCTGCTTATTAATTTGACTGGAATTAATTAAT	732
DB	670	TATGCTTGCTGCAATTTTTTTAGTGTGCTGCTTATTAATTTGACTGGAATTAATTAAT	729
QY	733	TACAGTTTTGATATGTGCCAAATAACTGCTCAGCGCAGAGAGTGTAAAGATCAGTAAT	792
DB	730	TACAGTTTTGATATGTGCCAAATAACTGCTCAGCGCAGAGAGTGTAAAGATCAGTAAT	789
QY	793	AGCAGCGAACTGTTGAATGTGAATGTTCTGAAACCTGGAAGGTGAAGCATGTGACATT	852
DB	790	AGCAGCGAATCTGTTGAATGTGAATGTTCTGAAACCTGGAAGGTGAAGCATGTGACATT	849
QY	853	CCTCACTGTACAGCAACTGTTGGTTTTCCTCATCGGCACTCTGCATTTCAAGTGTGTC	912
DB	850	CCTCACTGTACAGCAACTGTTGGTTTTCCTCATCGGCACTCTGCATTTCAAGTGTGTC	909
QY	913	AGAGGATGCTCTGCTTCTCAGACTGGCAGGGTCTCTGATGTTTCAGTTCCTGTACCACT	972
DB	910	AGAGGATGCTCTGCTTCTCAGACTGGCAGGGTCTCTGATGTTTCAGTTCCTGTACCACT	969

QY	973	AACCACTGCTTTTGGACTCGAGAGGAATATTCTAACTTTAAAGCTCCCGAGAGCATCTCAT	1032
DB	970	AACCACTGCTTTTGGACTCGAGAGGAATATTCTAACTTTAAAGCTCCCGAGAGCATCTCAT	1029
QY	1033	AAAGCTGTGCTCAATGGAAACATTATGTGGTGTGGAGGATATATCTCAACCACTCA	1092
DB	1030	AAAGCTGTGCTCAATGGAAACATTATGTGGTGTGGAGGATATATCTCAACCACTCA	1089
QY	1093	GATTATAACATGCTTAGCGCTATGACCTTGTCTTAGGGAGTGGCTTCCCAACCGT	1152
DB	1090	GATTATAACATGCTTAGCGCTATGACCTTGTCTTAGGGAGTGGCTTCCCAACCGT	1149
QY	1153	TCGTGAAACAATGTGGTGTGTAGATATGCTCATCTTTTGGCAATATACAAGGATAAAAT	1212
DB	1150	TCGTGAAACAATGTGGTGTGTAGATATGCTCATCTTTTGGCAATATACAAGGATAAAAT	1209
QY	1213	TACATGTATGGAGGAAATTTGATCCAATGGGAATGTGACCAATGAGTTGAGAGT	1272
DB	1210	TACATGTATGGAGGAAATTTGATCCAATGGGAATGTGACCAATGAGTTGAGAGT	1269
QY	1273	CACATTCATATGATGCTATGGGTGTGTGACCCCTAAAGGCAAGAGAGATGACGTG	1332
DB	1270	CACATTCATATGATGCTATGGGTGTGTGACCCCTAAAGGCAAGAGAGATGACGTG	1329
QY	1333	GTGGGCACTGTCACACATTTTACACTGAAGATGCCGAGTGTCTCATCTGCTCATC	1392
DB	1330	GTGGGCACTGTCACACATTTTACACTGAAGATGCCGAGTGTCTCATCTGCTCATC	1389
QY	1393	TTTGTGCTACTGCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAG	1452
DB	1390	TTTGTGCTACTGCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAG	1449
QY	1453	AAACATGAGGATATATACACACCCAGGCTGCCCTTGTGCAAGGGGTTACGGCCATAGC	1512
DB	1450	AAACATGAGGATATATACACACCCAGGCTGCCCTTGTGCAAGGGGTTACGGCCATAGC	1509
QY	1513	AGTGTTTACGACCATAGGACAGGCGCTTATACGTTTCATGTGGCTACAAGGCTTTTCACT	1572
DB	1510	AGTGTTTACGACCATAGGACAGGCGCTTATACGTTTCATGTGGCTACAAGGCTTTTCACT	1569
QY	1573	GCCAAATAGTACCGCTTGCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGG	1632
DB	1570	GCCAAATAGTACCGCTTGCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGG	1629
QY	1633	ACCATTTTAAGGACAGCGGATTTTTCGTTTACTTTGACACAGCTGTGATAGTGTGGA	1692
DB	1630	ACCATTTTAAGGACAGCGGATTTTTCGTTTACTTTGACACAGCTGTGATAGTGTGGA	1689
QY	1693	ACCATGCTGGTGTGTTGGGGGAAACACACAAATGACACATCTATGAGCCATGGGCGCAAA	1752
DB	1690	ACCATGCTGGTGTGTTGGGGGAAACACACAAATGACACATCTATGAGCCATGGGCGCAAA	1749
QY	1753	TGCTTTCTTTCAGATTTTATGCGCTTATGACATTTGCTGTCGCTGCTGCTGCTTCCC	1812
DB	1750	TGCTTTCTTTCAGATTTTATGCGCTTATGACATTTGCTGTCGCTGCTGCTGCTTCCC	1809
QY	1813	AGACTGTATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAAACAGC	1872
DB	1810	AGACTGTATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAAACAGC	1869
QY	1873	ACCATGTATGTTGCTGGTGTGTTCAATAGTCTCTCTCCTCAGGACATCTCTGTTTACCC	1932
DB	1870	ACCATGTATGTTGCTGGTGTGTTCAATAGTCTCTCTCCTCAGGACATCTCTGTTTACCC	1929
QY	1933	TCGGAACAGTGTGATGCGCATCGAGTCAAGCGCTTGTGTTAGCAGCAGGACCTGGTATT	1992
DB	1930	TCGGAACAGTGTGATGCGCATCGAGTCAAGCGCTTGTGTTAGCAGCAGGACCTGGTATT	1989
QY	1993	CGGTGTGTGTGAAACACAGGGTGTCTCAGTGTATCTGCTGGGCTGGCACTGATGAA	2052
DB	1990	CGGTGTGTGTGAAACACAGGGTGTCTCAGTGTATCTGCTGGGCTGGCACTGATGAA	2049
QY	2053	CAAGAAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTTGACCATGACAGATGT	2112



QY 4161 GTGAGTCTTGTGTGGCCAGCGCCCTGTGTGACATTTCTCAGCAGATGCCGATAGTGA 4220  
Db 1928 GTGAGTCTTGTGTGGCCAGCGCCCTGTGTGACATTTCTCAGCAGATGCCGATAGTGA 1987  
QY 4221 CAAGGAGAAGTCAGGAGCGGTGAGAAACCGGAAGCAGAGCCCTGTGACAGCTTGGGAC 4280  
Db 1988 CAAGGAGAAGTCAGGAGCGGTGAGAAACCGGAAGCAGAGCCCTGTGACAGCTTGGGAC 2047  
QY 4281 CTGCATCTGA 4290  
Db 2048 CTGCATCTGA 2057

RESULT 5  
US-09-893-238-1  
; Sequence 1, Application US/09893238  
; Patent No. US2002015093A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 8827  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-893-238-1

Query Match 6.8%; Score 291; DB 9; Length 8827;  
Best Local Similarity 100.0%; Pred. No. 4.7e-142;  
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 4000 GCCTTGGAAACAGATGAGAGGCTCTCTGATCTTATTTGGGGGAGTATAAGACTGTTCCC 4059  
Db 4266 GCCTTGGAAACAGATGAGAGGCTCTCTGATCTTATTTGGGGGAGTATAAGACTGTTCCC 4325  
QY 4060 AAACCATTTGACTGGAGCGGTGTTTGGCAACAAAGCGCTGCTCTCTGTTGTG 4119  
Db 4326 AAACCATTTGACTGGAGCGGTGTTTGGCAACAAAGCGCTGCTCTCTGTTGTG 4385  
QY 4120 AGGCTCCCTCGAGGCTGGGTGGATCCCTCTCTCTGGGAGTCAGGTCTTGTGTCGCC 4179  
Db 4386 AGGCTCCCTCGAGGCTGGGTGGATCCCTCTCTCTGGGAGTCAGGTCTTGTGTCGCC 4445  
QY 4180 AGCGCCCTGGTGGACATTTCTCAGCAGATCCGATGTTACAGGAGAGTACAGGAGCC 4239  
Db 4446 AGCGCCCTGGTGGACATTTCTCAGCAGATCCGATGTTACAGGAGAGTACAGGAGCC 4505  
QY 4240 GTGAGAAACCGGAAGCAGAGCCCTGTGACAGCTGGGACCTGCATCTGA 4290  
Db 4506 GTGAGAAACCGGAAGCAGAGCCCTGTGACAGCTGGGACCTGCATCTGA 4556

RESULT 6  
US-09-736-457-966/c  
; Sequence 966, Application US/09736457  
; Patent No. US20020168637A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary

; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aljun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 966  
; LENGTH: 246  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-736-457-966

Query Match 5.7%; Score 246; DB 9; Length 246;  
Best Local Similarity 100.0%; Pred. No. 2.1e-118;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2515 CTGCGAATAATGAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGC 2574  
Db 246 CTGCGAATAATGAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGC 187  
QY 2575 CTTCCGAAGATCAATGTCTCTACTGGTCTGGAGATATGTCCTCAATTTACAATAAGT 2634  
Db 186 CTTCCGAAGATCAATGTCTCTACTGGTCTGGAGATATGTCCTCAATTTACAATAAGT 127  
QY 2635 TTACTACAGTGGATCCGCTCAGCCCCAGTGTCTGGATTCTGGAATTTATCAGAA 2694  
Db 126 TTACTACAGTGGATCCGCTCAGCCCCAGTGTCTGGATTCTGGAATTTATCAGAA 67  
QY 2695 CCAGTACTCGGGAGTGAAGCTGCACCTGCATCAACCCACTCAATGGTAGTCTGT 2754  
Db 66 CCAGTACTCGGGAGTGAAGCTGCACCTGCATCAACCCACTCAATGGTAGTCTGT 7  
QY 2755 GAAAGG 2760  
Db 6 GAAAGG 1

RESULT 7  
US-09-902-941-966/c  
; Sequence 966, Application US/09902941  
; Patent No. US20020172952A1  
; GENERAL INFORMATION:  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Watanabe, Yoshihiro  
; APPLICANT: Johnson, Jeffrey C.  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Marnerakis, Margarita  
; APPLICANT: Carter, Darrick  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Vedvick, Thomas S.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: McNgabb, Andria  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.478C17  
; CURRENT APPLICATION NUMBER: US/09/902,941  
; CURRENT FILING DATE: 2001-07-10  
; NUMBER OF SEQ ID NOS: 2002  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 966  
; LENGTH: 246  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-902-941-966

Query Match  
Best Local Similarity 5.7%; Score 246; DB 9; Length 246;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGACATCTCAGAGCATGTCACCTTAACCCCATGGTGGC 2574  
Db 246 CTGCGAATAATGACATCTCAGAGCATGTCACCTTAACCCCATGGTGGC 187

QY 2575 CTTCCGAAGATCAATGTCCTACTGCTGGGAGATATGCCCATTTACAAATAGT 2634  
Db 186 CTTCCGAAGATCAATGTCCTACTGCTGGGAGATATGCCCATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTGCATGCTGGAAATTTATCAGAA 2694  
Db 126 TTACTACAGTGGATCCGCTCTGAGCCAGTGCATGCTGGAAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754  
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7

QY 2755 GAAAGG 2760  
Db 6 GAAAGG 1

## RESULT 8

US-09-849-626-966/c

; Sequence 966, Application US/09849626

; Publication No. US20020197669A1

; GENERAL INFORMATION:

; APPLICANT: Bangur, Chaitanya

; APPLICANT: Fanger, Gary

; APPLICANT: Wang, Aijun

; APPLICANT: Wang, Tongtong

; APPLICANT: Switzer, Anne

; APPLICANT: McNeill, Patricia

; APPLICANT: Clapper, Jonathan

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

; FILE REFERENCE: 210121.478C16

; CURRENT APPLICATION NUMBER: US/09/849,626

; CURRENT FILING DATE: 2001-05-03

; NUMBER OF SEQ ID NOS: 1926

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 966

; LENGTH: 246

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-849-626-966

Query Match  
Best Local Similarity 5.7%; Score 246; DB 9; Length 246;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGACATCTCAGAGCATGTCACCTTAACCCCATGGTGGC 2574  
Db 246 CTGCGAATAATGACATCTCAGAGCATGTCACCTTAACCCCATGGTGGC 187

QY 2575 CTTCCGAAGATCAATGTCCTACTGCTGGGAGATATGCCCATTTACAAATAGT 2634  
Db 186 CTTCCGAAGATCAATGTCCTACTGCTGGGAGATATGCCCATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTGCATGCTGGAAATTTATCAGAA 2694  
Db 126 TTACTACAGTGGATCCGCTCTGAGCCAGTGCATGCTGGAAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754  
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7

QY 2755 GAAAGG 2760  
Db 6 GAAAGG 1

## RESULT 9

US-10-017-754-966/c

; Sequence 966, Application US/10017754

; Publication No. US20030054363A1

; GENERAL INFORMATION:

; APPLICANT: Henderson, Robert A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Marnarakis, Margarita

; APPLICANT: Carter, Darrick

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.478C18

; CURRENT APPLICATION NUMBER: US/10/017,754

; CURRENT FILING DATE: 2001-10-29

; NUMBER OF SEQ ID NOS: 2004

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 966

; LENGTH: 246

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-017-754-966

Query Match  
Best Local Similarity 5.7%; Score 246; DB 14; Length 246;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGACATCTCAGAGCATGTCACCTTAACCCCATGGTGGC 2574  
Db 246 CTGCGAATAATGACATCTCAGAGCATGTCACCTTAACCCCATGGTGGC 187

QY 2575 CTTCCGAAGATCAATGTCCTACTGCTGGGAGATATGCCCATTTACAAATAGT 2634  
Db 186 CTTCCGAAGATCAATGTCCTACTGCTGGGAGATATGCCCATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATCCGCTCTGAGCCAGTGCATGCTGGAAATTTATCAGAA 2694  
Db 126 TTACTACAGTGGATCCGCTCTGAGCCAGTGCATGCTGGAAATTTATCAGAA 67

QY 2695 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754  
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7

QY 2755 GAAAGG 2760  
Db 6 GAAAGG 1

## RESULT 10

US-10-113-872-966/c

; Sequence 966, Application US/10113872

; Publication No. US20030170255A1

; GENERAL INFORMATION:

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Henderson, Robert A.

; APPLICANT: Kalos, Michael D.

; APPLICANT: Sleath, Paul R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Carter, Darrick

; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.478C19

; CURRENT APPLICATION NUMBER: US/10/113,872

; CURRENT FILING DATE: 2002-03-28

; NUMBER OF SEQ ID NOS: 2011  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 966  
; LENGTH: 246  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-113-872-966

Query Match  
Best Local Similarity 100.0%; Score 246; DB 14; Length 246;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATATGAGTCATCTCAGAGCATGTCCAAAGCTCAGCTTAAACCCATGGTGGC 2574  
Db CTGCGAATATGAGTCATCTCAGAGCATGTCCAAAGCTCAGCTTAAACCCATGGTGGC 187

QY 2575 CTTCCGAAGATCAATGTCTTACTGCTGCTGGAGATATGTCCTCCATTTACAATAGT 2634  
Db CTTCCGAAGATCAATGTCTTACTGCTGCTGGAGATATGTCCTCCATTTACAATAGT 127

QY 2635 TTACTACAGTGGATGCGCTCTGAGCCAGTGTGCTGGATTTCTGGAATTTTATCAGAA 2694  
Db TTACTACAGTGGATGCGCTCTGAGCCAGTGTGCTGGATTTCTGGAATTTTATCAGAA 67

QY 2695 CCCAGTACTCGGGAGTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 2754  
Db CCCAGTACTCGGGAGTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTGT 7

QY 2755 GAAAGG 2760  
Db GAAAGG 1

RESULT 11  
US-10-277-216-5  
; Sequence 5, Application US/10277216  
; Publication No. US2004002470A1  
; GENERAL INFORMATION:  
; APPLICANT: KEITH, TIM  
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,  
; FILE REFERENCE: 2976-4051  
; CURRENT APPLICATION NUMBER: US/10/277,216  
; PRIOR FILING DATE: 2002-10-17  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 09/834,597  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 09/548,797  
; PRIOR FILING DATE: 2000-04-13  
; NUMBER OF SEQ ID NOS: 420  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 207433  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-277-216-5

Query Match  
Best Local Similarity 100.0%; Score 234; DB 15; Length 207433;  
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2951 CTGAAAATTTTCAGGCTACTGCTAGTCAATGCTGGAGCAACAGGCTGTGGCT 3010  
Db CTGAAAATTTTCAGGCTACTGCTAGTCAATGCTGGAGCAACAGGCTGTGGCT 9872

QY 3011 GGTGTACTGATCCAGCAATCTGCAAGGGAATGATAGAGGTTCTTATAAGGAC 3070  
Db GGTGTACTGATCCAGCAATCTGCAAGGGAATGATAGAGGTTCTTATAAGGAC 9932

QY 3071 CAGTGAAGATGCTTCCGAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAAT 3130  
Db CAGTGAAGATGCTTCCGAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAAT 9992

QY 3131 CCAGCATGTCTAGAGGACAGCAGATACAACTGGTCTTTCACTGCTCCAG 3184  
Db CCAGCATGTCTAGAGGACAGCAGATACAACTGGTCTTTCACTGCTCCAG 10046

RESULT 13  
US-10-198-846-5869/c  
; Sequence 5869, Application US/10198846  
; Publication No. US2003009997A1  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Lillie, James  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steinmann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; FILE REFERENCE: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; FILE REFERENCE: THERAPY OF BREAST CANCER  
; CURRENT APPLICATION NUMBER: US/10/198,846  
; PRIOR FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5869  
; LENGTH: 625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-198-846-5869/c

QY 3131 CCAGCATGTCTAGAGGACAGCAGATACAACTGGTCTTTCACTGCTCCAG 3184  
Db 9993 CCAGCATGTCTAGAGGACAGCAGATACAACTGGTCTTTCACTGCTCCAG 10046

RESULT 12  
US-10-126-022-5  
; Sequence 5, Application US/10126022  
; Publication No. US20040023215A1  
; GENERAL INFORMATION:  
; APPLICANT: KEITH, TIM  
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,  
; FILE REFERENCE: OBESITY, AND INFLAMMATORY BOWEL DISEASE  
; FILE REFERENCE: 2976-4039U92  
; CURRENT APPLICATION NUMBER: US/10/126,022  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 09/834,597  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 09/548,797  
; PRIOR FILING DATE: 2000-04-13  
; NUMBER OF SEQ ID NOS: 420  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 207433  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-126-022-5

Query Match  
Best Local Similarity 100.0%; Score 234; DB 16; Length 207433;  
Matches 234; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2951 CTGAAAATTTTCAGGCTACTGCTAGTCAATGCTGGAGCAACAGGCTGTGGCT 3010  
Db 9813 CTGAAAATTTTCAGGCTACTGCTAGTCAATGCTGGAGCAACAGGCTGTGGCT 9872

QY 3011 GGTGTACTGATCCAGCAATCTGCAAGGGAATGATAGAGGTTCTTATAAGGAC 3070  
Db GGTGTACTGATCCAGCAATCTGCAAGGGAATGATAGAGGTTCTTATAAGGAC 9932

QY 3071 CAGTGAAGATGCTTCCGAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAAT 3130  
Db CAGTGAAGATGCTTCCGAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAAT 9992

QY 3131 CCAGCATGTCTAGAGGACAGCAGATACAACTGGTCTTTCACTGCTCCAG 3184  
Db CCAGCATGTCTAGAGGACAGCAGATACAACTGGTCTTTCACTGCTCCAG 10046



```

;
; NAME/KEY: misc_feature
; LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5869

Query Match
5.3%; Score 228; DB 14; Length 625;
Best Local Similarity 99.6%; Pred. No. 6.8e-109;
Matches 278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2699 GTACTCGGAGCTGAAGGCTGCAACTGTCATCAACCCCACTCAATGGTAGTGTCTGTGAAA 2758
DB 340 GTACTCGGAGCTGAAGGCTGCAACTGTCATCAACCCCACTCAATGGTAGTGTCTGTGAAA 281
QY 2759 GGCCTGCAAAACACAGCTGTAAGCAGTGCAGGACACCATGTGCCCTTGAGGACAGCATGTG 2818
DB 280 GGCCTGCAAAACACAGCTGTAAGCAGTGCAGGACACCATGTGCCCTTGAGGACAGCATGTG 221
QY 2819 GAGATTGCACCGCCGACCTCTGAGTGCATGTGGTGACGACCAATGAAGCAGTGTGTGG 2878
DB 220 GAGATTGCACCGCCGACCTCTGAGTGCATGTGGTGACGACCAATGAAGCAGTGTGTGG 161
QY 2879 ACTCCAATCCATATGTGGCTCTCTCCCTTTGGCCAGTGTATGAAATGGTATACATGA 2938
DB 160 ACTCCAATCCATATGTGGCTCTCTCCCTTTGGCCAGTGTATGAAATGGTATACATGA 101
QY 2939 GCACCTGCCCTGCAAAATTTGTCAGGCTACTGTACCT 2977
DB 100 GCACCTGCCCTGCAAAATTTGTCAGGCTACTGTACCT 62

RESULT 14
US-09-864-761-15900
; Sequence 15900, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aesomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

;
; NAME/KEY: misc_feature
; LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29
; OTHER INFORMATION: n = A,T,C or G
US-09-864-761-15900

Query Match
4.9%; Score 210; DB 9; Length 531;
Best Local Similarity 100.0%; Pred. No. 2.1e-99;
Matches 210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1238 CAACTGGGAATGTGACCAATGAGTGTGAGAGTGTTCACATTCAATAAGTCAATGGGTGT 1297
DB 184 CAACTGGGAATGTGACCAATGAGTGTGAGAGTGTTCACATTCAATAAGTCAATGGGTGT 243
QY 1298 TGTGACCCCTAAGGCAAGGACGATGTCAGTGTGGGCACTCTGCACACATTGTGA 1357
DB 244 TGTGACCCCTAAGGCAAGGACGATGTCAGTGTGGGCACTCTGCACACATTGTGA 303
QY 1358 CACTGAAGAATGCGCGAGTGTGATGTCATCTTTGGTCACTGCCCTCTCTATGGAT 1417
DB 304 CACTGAAGAATGCGCGAGTGTGATGTCATCTTTGGTCACTGCCCTCTCTATGGAT 363
QY 1418 ATATAAGCAATGTGCAGGAATATGATTGG 1447
DB 364 ATATAAGCAATGTGCAGGAATATGATTGG 393

RESULT 15
US-09-864-761-32404
; Sequence 32404, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aesomica-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 32404
; LENGTH: 193
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132773.14
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
; OTHER INFORMATION: EST HUMAN HIT: BF313142.1, EVALUATE 1.00e-104
; OTHER INFORMATION: NT HIT: AF218896.1, EVALUATE 1.00e-105
; OTHER INFORMATION: SWISSPROT HIT: Q19981, EVALUATE 1.00e-05
US-09-864-761-32404

Query Match 4.5%; Score 193; DB 9; Length 193;
Best Local Similarity 100.0%; Pred. No. 1.8e-90;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1245 GAATGTGACCAATGAGTTCAGAGTGTTCACATTCATATGAGTCATGGTGTGTTGAC 1304
Db 1 GAATGTGACCAATGAGTTCAGAGTGTTCACATTCATATGAGTCATGGTGTGTTGAC 60

QY 1305 CCTAAGGCAAGGAGCAGTATGAGTGTTCAGTGTGGGCACTCTGCACACATTTTACACTGAA 1364
Db 61 CCTAAGGCAAGGAGCAGTATGAGTGTGGGCACTCTGCACACATTTTACACTGAA 120

QY 1365 GAATGCCGAGTGTGATGCTGTGTCATCTTTGGTGCATGCGCTCTCTATGATATATAAG 1424
Db 121 GAATGCCGAGTGTGATGCTGTGTCATCTTTGGTGCATGCGCTCTCTATGATATATAAG 180

QY 1425 CAATGTGCAGGAA 1437
Db 181 CAATGTGCAGGAA 193

RESULT 16
US-09-864-761-21460/c
; Sequence 21460, Application US/09864761
; Patent No US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
```

```

; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21460
; LENGTH: 182
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132773.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
; OTHER INFORMATION: NT HIT: AF218902.1, EVALUATE 2.00e-98
; OTHER INFORMATION: EST HUMAN HIT: AI372810.1, EVALUATE 2.00e-98
; OTHER INFORMATION: SWISSPROT HIT: P22897, EVALUATE 8.00e-03
US-09-864-761-21460

Query Match 4.2%; Score 182; DB 9; Length 182;
Best Local Similarity 100.0%; Pred. No. 1.2e-84;
Matches 182; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2545 TCCAGCTCACCTTAACCCCATGGGTGGCCCTTCGGAAGATCAATGTGTCTTACTGGTGC 2604
Db 182 TCCAGCTCACCTTAACCCCATGGGTGGCCCTTCGGAAGATCAATGTGTCTTACTGGTGC 123

QY 2605 TGGGAAGATATGTCCCATTTTACAATAGTTTACTACAGTGCCTGCTGAGCCCACT 2664
Db 122 TGGGAAGATATGTCCCATTTTACAATAGTTTACTACAGTGCCTGCTGAGCCCACT 63

QY 2665 GATGCTGGATTCGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGCTCAACC 2724
Db 62 GATGCTGGATTCGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGCTCAACC 3

QY 2725 TG 2726
Db 2 TG 1
```

RESULT 17  
US-09-864-761-4720/c  
; Sequence 4720, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 4720  
; LENGTH: 375  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL13273.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.6  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1  
US-09-864-761-4720

Query Match 4.0%; Score 173; DB 9; Length 375;

Best Local Similarity 100.0%; Pred. No. 6.7e-80;  
Matches 173; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2592 GTCTACTGCTGCTGGAGATATCTCCCAATTTTACAAATAGTTTACTACATGATGCC 2651  
DB 375 GTCCCTACTGCTGCTGGAGATATCTCCCAATTTTACAAATAGTTTACTACATGATGCC 316  
QY 2652 GTCTAGCCCCAGTGTCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTTCGGGGACT 2711  
DB 315 GTCTAGCCCCAGTGTCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTTCGGGGACT 256  
QY 2712 GAAGGTGCAACTGATCAATCAACCCACTCAATGCTGTCTGTGAAGGCCTG 2764  
DB 255 GAAGGTGCAACTGATCAATCAACCCACTCAATGCTGTCTGTGAAGGCCTG 203

RESULT 18  
US-09-864-761-3681/c  
; Sequence 3681, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecomica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 3681  
; LENGTH: 402  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:



```

; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7610
; LENGTH: 212
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (140)
; OTHER INFORMATION: n=A, T, C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (142)
; OTHER INFORMATION: n=A, T, C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (162)
; OTHER INFORMATION: n=A, T, C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (196)
; OTHER INFORMATION: n=A, T, C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (204)
; OTHER INFORMATION: n=A, T, C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (205)
; OTHER INFORMATION: n=A, T, C or G
; US-10-040-862-7610

```

	Query Match	3.2%	Score 139;	DB 14;	Length 212;
	Best Local Similarity	100.0%;	Prod. No. 5.5e-62;		
	Matches 139;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2913	CCAGTGATGGAATGGTATACGATGACACCTGCCCCCTGAAATGTTTCAGGCTACTG	2972		
DB	1	CCAGTGATGGAATGGTATACGATGACACCTGCCCCCTGAAATGTTTCAGGCTACTG	60		
QY	2973	TACCTGTAGTCATTGCTTGGAGCAACACAGGCTGTGGCTGGTGTACTGATCCCAAGCAATAC	3032		
DB	61	TACCTGTAGTCATTGCTTGGAGCAACACAGGCTGTGGCTGGTGTACTGATCCCAAGCAATAC	120		
QY	3033	TGGCAAGGGAAATGCATA	3051		
DB	121	TGGCAAGGGAAATGCATA	139		

RESULT 21  
US-10-057-475B-7610  
US-10-057-475B-7610 ; Sequence 7610, Application US/10057475B  
; Publication No. US20040002068A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Clapper, Jonathan David  
; APPLICANT: Wang, Aijun  
; APPLICANT: Ordonez, Nadia  
; APPLICANT: Carter, Lauren  
; APPLICANT: McNeill, Patricia Dianne  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; TITLE OF INVENTION: Hematological Malignancies  
; FILE REFERENCE: 014058-014402US  
; CURRENT APPLICATION NUMBER: US/10/057.475B

```

CURRENT FILING DATE: 2002-01-22
PRIOR APPLICATION NUMBER: US 60/186,126
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: US 60/190,479
PRIOR FILING DATE: 2000-03-17
PRIOR APPLICATION NUMBER: US 60/200,545
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: US 60/200,303
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,779
PRIOR FILING DATE: 2000-04-28
PRIOR APPLICATION NUMBER: US 60/200,999
PRIOR FILING DATE: 2000-05-01
PRIOR APPLICATION NUMBER: US 60/202,084
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 60/206,201
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: US 60/218,950
PRIOR FILING DATE: 2000-07-14
PRIOR APPLICATION NUMBER: US 60/222,903
PRIOR FILING DATE: 2000-08-03
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 10979
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7610
LENGTH: 212
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(212)
OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-7610

```

	Query Match	3.2%;	Score 139;	DB 15;	Length 212;
	Best Local Similarity	100.0%;	Pred. No. 5.5e-62;		
	Matches 139;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	2913	CCAGCTGTATGGAAATGTTACGATGAGCACCCTGCCCCCTGAAAAATTGTTCCAGGCTACTG	2972		
Db	1	CCAGTGTATGGAAATGTTACGATGAGCACCCTGCCCCCTGAAAAATTGTTCCAGGCTACTG	60		
Qy	2973	TACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGTACTGTATCCAGCAATAC	3032		
Db	61	TACCTGTAGTCATTGCTTGGAGCAACCAGGCTGTGGCTGTACTGTATCCAGCAATAC	120		
Qy	3033	TGGCAAGGGAATGCATA	3051		
Db	121	TGGCAAGGGAATGCATA	139		

RESULT 22  
US-10-154-884B-7610  
; Sequence 7610, Application US/10154884B  
; Publication No. US20040005561A1  
; GENERAL INFORMATION:  
; APPLICANT: Gaiger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; APPLICANT: Retter, Marc W.  
; APPLICANT: Corixa Corporation  
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy  
; TITLE OF INVENTION: Hematological Malignancies  
; FILE REFERENCE: 014058-013521US  
; CURRENT APPLICATION NUMBER: US/10/154,884B  
; CURRENT FILING DATE: 2002-05-23  
; PRIOR APPLICATION NUMBER: US 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: US 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: US 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: US 60/200,303

```

; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7610
; LENGTH: 212
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(212)
; OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-7610

Query Match          3.2%; Score 139; DB 15; Length 212;
Best Local Similarity 100.0%; Pred. No. 5.5e-62;
Matches 139; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2913 CCAGTGTATGAATGGTATACGATGAGCAGCCTGCCCTGAAATTTCTCAGGCTACTG 2972
Db 1 CCAAGTGTATGAATGGTATACGATGAGCAGCCTGCCCTGAAATTTCTCAGGCTACTG 60

QY 2973 TACCTGTAGTCAATGCTTGAGCAACACGAGCTGTGGCTGGTGTACTGTATCCAGCAATAC 3032
Db 61 TACCTGTAGTCAATGCTTGAGCAACACGAGCTGTGGCTGGTGTACTGTATCCAGCAATAC 120

QY 3033 TGGCAAGGGAATGCATA 3051
Db 121 TGGCAAGGGAATGCATA 139

RESULT 23
US-09-864-761-20449/c
; Sequence 20449, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aomics-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 20449
; LENGTH: 137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132773.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EST HUMAN HIT: BF672370.1, EVALUATE 2.00e-71
; OTHER INFORMATION: NT HIT: AB011120.1, EVALUATE 1.00e-71
; OTHER INFORMATION: SWISSPROT HIT: Q19981, EVALUATE 3.00e-11
US-09-864-761-20449

Query Match          3.2%; Score 137; DB 9; Length 137;
Best Local Similarity 100.0%; Pred. No. 6.1e-61;
Matches 137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3185 CTTGCCAATGCAACGCCACAGTAAATGCATCAATCAGACGATCTGTGAGAAGTGTGAGA 3244
Db 137 CTTGCCAATGCAACGCCACAGTAAATGCATCAATCAGACGATCTGTGAGAAGTGTGAGA 78

QY 3245 ACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3304
Db 77 ACCTGACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 18

QY 3305 ATGCGAGGGAATGTCTAG 3321
Db 17 ATGCGAGGGAATGTCTAG 1

RESULT 24
US-10-027-632-264639
; Sequence 264639, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12

```

;; PRIOR APPLICATION NUMBER: US 60/198,676  
;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: US 60/193,483  
;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: US 60/185,218  
;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: US 60/167,363  
;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358  
;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002  
;; PRIOR FILING DATE: 1999-08-09  
;; NUMBER OF SEQ ID NOS: 325720  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 264639  
;; LENGTH: 594  
;; TYPE: DNA  
;; ORGANISM: Human  
US-10-027-632-264639

Query Match 2.0%; Score 86; DB 15; Length 594;  
Best Local Similarity 100.0%; Pred. No. 5.1e-34;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1786 GCCTGTGACCGTGTGCTGCTTCCAGACCTGATCTCCACCATGATGTCAACAGATTT 1845  
DB 144 GCCTGTGACCGTGTGCTGCTTCCAGACCTGATCTCCACCATGATGTCAACAGATTT 203

QY 1846 GCCATTTCAGAGTCTTACAAACAG 1871  
DB 204 GCCATTTCAGAGTCTTACAAACAG 229

RESULT 25  
US-10-027-632-264640  
;; Sequence 264640, Application US/10027632  
;; Publication No. US20030204075A9  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.  
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

;; FILE REFERENCE: 108827.129  
;; CURRENT APPLICATION NUMBER: US/10/027,632  
;; PRIOR FILING DATE: 2002-04-30  
;; PRIOR APPLICATION NUMBER: US 60/218,006  
;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/198,676  
;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: US 60/193,483  
;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: US 60/185,218  
;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: US 60/167,363  
;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358  
;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002  
;; NUMBER OF SEQ ID NOS: 325720  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 264640  
;; LENGTH: 594  
;; TYPE: DNA  
;; ORGANISM: Human  
US-10-027-632-264640

Query Match 2.0%; Score 86; DB 15; Length 594;  
Best Local Similarity 100.0%; Pred. No. 5.1e-34;  
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1786 GCCTGTGACCGTGTGCTGCTTCCAGACCTGATCTCCACCATGATGTCAACAGATTT 1845  
DB 144 GCCTGTGACCGTGTGCTGCTTCCAGACCTGATCTCCACCATGATGTCAACAGATTT 203

QY 1846 GGCCATTTCAGAGTCTTACAAACAG 1871  
DB 204 GGCCATTTCAGAGTCTTACAAACAG 229

RESULT 26  
US-10-027-632-270240  
;; Sequence 270240, Application US/10027632  
;; Publication No. US20030204075A9  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.  
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
;; FILE REFERENCE: 108827.129  
;; CURRENT APPLICATION NUMBER: US/10/027,632  
;; PRIOR FILING DATE: 2002-04-30  
;; PRIOR APPLICATION NUMBER: US 60/218,006  
;; PRIOR FILING DATE: 2000-07-12  
;; PRIOR APPLICATION NUMBER: US 60/198,676  
;; PRIOR FILING DATE: 2000-04-20  
;; PRIOR APPLICATION NUMBER: US 60/193,483  
;; PRIOR FILING DATE: 2000-03-29  
;; PRIOR APPLICATION NUMBER: US 60/185,218  
;; PRIOR FILING DATE: 2000-02-24  
;; PRIOR APPLICATION NUMBER: US 60/167,363  
;; PRIOR FILING DATE: 1999-11-23  
;; PRIOR APPLICATION NUMBER: US 60/156,358  
;; PRIOR FILING DATE: 1999-09-28  
;; PRIOR APPLICATION NUMBER: US 60/146,002  
;; NUMBER OF SEQ ID NOS: 325720  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 270240  
;; LENGTH: 629  
;; TYPE: DNA  
;; ORGANISM: Human  
US-10-027-632-270240

Query Match 1.8%; Score 79; DB 15; Length 629;  
Best Local Similarity 100.0%; Pred. No. 2.5e-30;  
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3565 CAAAACAGGATTTGGACATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACATCACC 3624  
DB 347 CAAAACAGGATTTGGACATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACATCACC 406

QY 3625 TGGGCTGCCAGTTTCTCAG 3643  
DB 407 TGGGCTGCCAGTTTCTCAG 425

RESULT 27  
US-09-908-975-15421  
;; Sequence 15421, Application US/09908975  
;; Publication No. US20030165843A1  
;; GENERAL INFORMATION:  
;; APPLICANT: SHOSHAN, Avi  
;; APPLICANT: WASSERMAN, Alon  
;; APPLICANT: MINTZ, Eli  
;; APPLICANT: FAIGLER, Simchon  
;; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE  
;; FILE REFERENCE: 36688-0005  
;; CURRENT APPLICATION NUMBER: US/09/908,975  
;; PRIOR FILING DATE: 2001-07-20  
;; PRIOR APPLICATION NUMBER: US 60/287,724  
;; PRIOR FILING DATE: 2001-05-02  
;; PRIOR APPLICATION NUMBER: US 60/221,607  
;; PRIOR FILING DATE: 2000-07-28  
;; NUMBER OF SEQ ID NOS: 32337  
;; SOFTWARE: PatentIn version 3.0

```
; SEQ ID NO 15421
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-15421

Query Match
Best Local Similarity 1.4%; Score 60; DB 10; Length 60;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3547 GTGGTACTCTCTGACGAAACAAACAGGAGTTTGGACATGTTTCATCAATGCCCTCCAAGAAAT 3606
Db 1 GTGGTACTCTCTGACGAAACAAACAGGAGTTTGGACATGTTTCATCAATGCCCTCCAAGAAAT 60

RESULT 28
US-09-893-238-5
; Sequence 5, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1998-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-5

Query Match
Best Local Similarity 1.3%; Score 55; DB 9; Length 90050;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3130 TCCAGCATGTGCTAGAGACAGACAGATACAACTGGTCTTTTCATTCACCTGTCCAG 3184
Db 3184 TCCAGCATGTGCTAGAGACAGACAGATACAACTGGTCTTTTCATTCACCTGTCCAG 31894

RESULT 29
US-09-893-238-10
; Sequence 10, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1998-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
```

```
; ORGANISM: Mus musculus
US-09-893-238-10

Query Match
Best Local Similarity 1.2%; Score 53; DB 9; Length 1051;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 GGCTGGTGGCGGAGCAATGCCAGCACTGGGGGGCCGCTTCAGACTAACTGG 419
Db 571 GGCTGGTGGCGGAGCAATGCCAGCACTGGGGGGCCGCTTCAGACTAACTGG 623

RESULT 30
US-09-893-238-8
; Sequence 8, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1998-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2419
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-8

Query Match
Best Local Similarity 1.1%; Score 47; DB 9; Length 2419;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1624 CAGATCTGGACCATCTTAAAGGACAGCCGATTTTCCGTTACTTGA 1670
Db 1890 CAGATCTGGACCATCTTAAAGGACAGCCGATTTTCCGTTACTTGA 1936

RESULT 31
US-10-027-632-108988/c
; Sequence 108988, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
```



```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 108988
; LENGTH: 468
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(468)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-108988

Query Match
Best Local Similarity 0.9%; Score 38; DB 15; Length 468;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3417 ATGTGAGCTAGAAATCGATACCAAGGAAACCCCTCTCA 3454
DB 414 ATGTGAGCTAGAAATCGATACCAAGGAAACCCCTCTCA 377

RESULT 32
US-09-893-238-4
; Sequence 4, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 5973
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-4

Query Match
Best Local Similarity 0.8%; Score 35; DB 9; Length 5973;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 547 TGGGACCAATTTATATGTTTATGATGGGACTCAAT 581
DB 4379 TGGGACCAATTTATATGTTTATGATGGGACTCAAT 4413

RESULT 33
US-09-908-975-29990
; Sequence 29990, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchaon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
```

```
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29990
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-29990

Query Match
Best Local Similarity 0.7%; Score 32; DB 10; Length 65;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2860 AACATGAAGCAGTGTGTGGACTCCAATGCTA 2891
DB 34 AACATGAAGCAGTGTGTGGACTCCAATGCTA 65

RESULT 34
US-09-893-238-3/c
; Sequence 3, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 17056
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-3

Query Match
Best Local Similarity 0.7%; Score 32; DB 9; Length 17056;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 844 TGTGACATTCTCCTCAGTACAGACAAGTGTGG 875
DB 3978 TGTGACATTCTCCTCAGTACAGACAAGTGTGG 3947

RESULT 35
US-09-893-238-23/c
; Sequence 23, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 23
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-09-893-238-23

Query Match      0.7%; Score 29; DB 9; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 552 CCATTATATGTTTATGATGGGACTCAA 580
Db 29 CCATTATATGTTTATGATGGGACTCAA 1

RESULT 36
US-10-419-296-1
; Sequence 1, Application US/10419296
; Publication No. US20040014141A1
; GENERAL INFORMATION:
; APPLICANT: Woolf, Clifford J.
; APPLICANT: Samad, Tarek A.
; TITLE OF INVENTION: DRG11-RESPONSIVE (DRAGON) GENE FAMILY
; FILE REFERENCE: 00786/419002
; CURRENT APPLICATION NUMBER: US/10/419,296
; PRIOR FILING DATE: 2003-04-17
; PRIOR APPLICATION NUMBER: 60/373,519
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1332
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-419-296-1

Query Match      0.7%; Score 29; DB 15; Length 1332;
Best Local Similarity 100.0%; Pred. No. 0.0006;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 CCCGCGCGCGCTGTGCTGCTGCTGCTG 232
Db 114 CCCGCGCGCGCTGTGCTGCTGCTGCTG 142

RESULT 37
US-10-029-386-22666
; Sequence 22666, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22666
; LENGTH: 1622
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL133445.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
; OTHER INFORMATION: SWISSPROT HIT: Q05175, EVALUATE 4.90e+00

; OTHER INFORMATION: NT HIT: g115595198, EVALUATE 1.60e+00
; OTHER INFORMATION: EST_HUMAN HIT: AA700990.1, EVALUATE 0.00e+00
US-10-029-386-22666

Query Match      0.7%; Score 29; DB 14; Length 1622;
Best Local Similarity 100.0%; Pred. No. 0.0006;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 240 GGCGGAGCGCGCGCGCGCGCGCGCGG 268
Db 967 GGCGGAGCGCGCGCGCGCGCGCGCGG 995

RESULT 38
US-10-311-455-169
; Sequence 169, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: Olek, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; TITLE OF INVENTION: cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 169
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-169

Query Match      0.6%; Score 27; DB 14; Length 5312;
Best Local Similarity 100.0%; Pred. No. 0.0072;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 249 CGCGCGCGCGCGCGCGCGCGGTGCGG 275
Db 5249 CGCGCGCGCGCGCGCGCGCGGTGCGG 5275

RESULT 39
US-09-893-238-48
; Sequence 48, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 24
; TYPE: DNA
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Primer
US-09-893-238-48

Query Match          0.6%; Score 24; DB 9; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.22;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1631 GGACCATTTCTTAAGCAGCGCGAT 1654
Db 1 GGACCATTTCTTAAGCAGCGCGAT 24

RESULT 40
US-09-783-590-7412
; Sequence 7412, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Hacdong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; PRIOR FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7412
; LENGTH: 498
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (141)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (152)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (271)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (319)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (325)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (332)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (333)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (337)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (357)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (362)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (380)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
```

```
; LOCATION: (384)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (388)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (390)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (393)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (394)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (395)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (397)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (398)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (416)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (420)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (438)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (444)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (448)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (452)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (455)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (463)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (466)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (468)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (471)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (483)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (491)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (493)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (497)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-783-590-7412
```

```
Query Match          0.6%; Score 24; DB 9; Length 498;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 547  
LENGTH: 1372  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)..(1372)  
OTHER INFORMATION: unsure at all n locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_100494C.1  
US-10-424-599-547

Query Match 0.6%; Score 24; DB 12; Length 1372;  
Best Local Similarity 100.0%; Pred. No. 0.26;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203  
DB 529 GCTGCTGCTGCTGCTGCTGCTGCT 506

RESULT 46  
US-10-424-599-21459/c  
Sequence 21459, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 21459  
LENGTH: 1756  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_119382C.1  
US-10-424-599-21459

Query Match 0.6%; Score 24; DB 12; Length 1756;  
Best Local Similarity 100.0%; Pred. No. 0.26;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203  
DB 1198 GCTGCTGCTGCTGCTGCTGCTGCT 1175

RESULT 47  
US-09-814-353-21391/c  
Sequence 21391, Application US/09814353  
Publication No. US20030165831A1  
GENERAL INFORMATION:  
APPLICANT: Lee, John  
APPLICANT: Thompson, Pamela  
APPLICANT: Lillie, James  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
IDENTIFICATION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
THERAPY OF OVARIAN CANCER  
FILE REFERENCE: MRI-006B  
CURRENT APPLICATION NUMBER: US/09/814,353

CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/191,031  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: US 60/207,124  
PRIOR FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: US 60/211,940  
PRIOR FILING DATE: 2000-06-15  
PRIOR APPLICATION NUMBER: US 60/216,820  
PRIOR FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: US 60/220,661  
PRIOR FILING DATE: 2000-07-25  
PRIOR APPLICATION NUMBER: US 60/257,672  
PRIOR FILING DATE: 2000-12-21  
NUMBER OF SEQ ID NOS: 22037  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 21391  
LENGTH: 2180  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1, 2, 45, 2179, 2180  
OTHER INFORMATION: n = A,T,C or G  
US-09-814-353-21391

Query Match 0.6%; Score 24; DB 10; Length 2180;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203  
DB 1050 GCTGCTGCTGCTGCTGCTGCTGCT 1027

RESULT 48  
US-10-437-171-3/c  
Sequence 3, Application US/10437171  
Publication No. US20030235564A1  
GENERAL INFORMATION:  
APPLICANT: Doll, Bruce  
APPLICANT: Fu, Huihua  
APPLICANT: Hollinger, Jeffrey O.  
APPLICANT: Sifer, Charles  
TITLE OF INVENTION: Compositions and Devices Comprising or Encoding the Run X2  
FILE REFERENCE: 1915/14014US02  
CURRENT APPLICATION NUMBER: US/10/437,171  
CURRENT FILING DATE: 2003-05-13  
PRIOR APPLICATION NUMBER: 60/380,554  
PRIOR FILING DATE: 2002-05-13  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 3  
LENGTH: 2294  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic Peptide  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (1)..(1644)  
US-10-437-171-3

Query Match 0.6%; Score 24; DB 15; Length 2294;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203  
DB 223 GCTGCTGCTGCTGCTGCTGCTGCT 200

RESULT 49

US-10-437-171-1/c  
; Sequence 1, Application US/10437171  
; Publication No. US20030235564A1  
; GENERAL INFORMATION:  
; APPLICANT: Doll, Bruce  
; APPLICANT: Fu, Huibua  
; APPLICANT: Hollinger, Jeffrey O.  
; APPLICANT: Sier, Charles  
; TITLE OF INVENTION: Compositions and Devices Comprising or Encoding the Run X2  
; TITLE OF INVENTION: Protein and Method of Use  
; FILE REFERENCE: 1915/14014US02  
; CURRENT APPLICATION NUMBER: US/10/437,171  
; CURRENT FILING DATE: 2003-05-13  
; PRIOR APPLICATION NUMBER: 60/380,554  
; PRIOR FILING DATE: 2002-05-13  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 1  
; LENGTH: 3334  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Primer  
US-10-437-171-1  
Query Match 0.6%; Score 24; DB 15; Length 3334;  
Best Local Similarity 100.0%; Pred. No. 0.27;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203  
DB 531 GCTGCTGCTGCTGCTGCTGCTGCT 508  
RESULT 50  
US-10-424-599-56004/c  
; Sequence 56004, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 56004  
; LENGTH: 4964  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_21582C.1  
US-10-424-599-56004  
Query Match 0.6%; Score 24; DB 12; Length 4964;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203  
DB 405 GCTGCTGCTGCTGCTGCTGCTGCT 362  
RESULT 51  
US-10-198-846-9854/c  
; Sequence 9854, Application US/10198846  
; Publication No. US2003009974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen  
; APPLICANT: Steinmann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-049  
; CURRENT APPLICATION NUMBER: US/10/198,846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9854  
; LENGTH: 5085  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1, 5070, 5071, 5072, 5073, 5074, 5075, 5076, 5077, 5078,  
; LOCATION: 5079, 5080, 5081, 5082, 5083, 5084, 5085  
; OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-9854

Query Match 0.6%; Score 24; DB 14; Length 5085;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203  
DB 1432 GCTGCTGCTGCTGCTGCTGCTGCT 1409

RESULT 52  
US-09-819-104A-6/c  
; Sequence 6, Application US/09819104A  
; Publication No. US20030027137A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, J. Don  
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: UMG-030  
; CURRENT APPLICATION NUMBER: US/09/819,104A  
; CURRENT FILING DATE: 2001-03-27  
; PRIOR APPLICATION NUMBER: 60/193,138  
; PRIOR FILING DATE: 2000-03-29  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 7386  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(7386)  
US-09-819-104A-6  
Query Match 0.6%; Score 24; DB 10; Length 7386;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203  
DB 1519 GCTGCTGCTGCTGCTGCTGCTGCT 1496

RESULT 53  
US-09-819-104A-3/c  
; Sequence 3, Application US/09819104A  
; Publication No. US20030027137A1  
; GENERAL INFORMATION:  
; APPLICANT: Chen, J. Don  
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR

```

; FILE REFERENCE: UMG-030
; CURRENT APPLICATION NUMBER: US/09/819,104A
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/193,138
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 7521
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(7521)
US-09-819-104A-3

Query Match      0.6%; Score 24; DB 10; Length 7521;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
Db 1504 GCTGCTGCTGCTGCTGCTGCTGCT 1481

RESULT 54
US-10-414-692-14/c
; Sequence 14, Application US/10414692
; Publication No. US20030228607A1
; GENERAL INFORMATION:
; APPLICANT: X-Ceptor Therapeutics, Inc.
; TITLE OF INVENTION: Screening method and modulators having an improved therapeutic
; FILE REFERENCE: 8012-002-US
; CURRENT APPLICATION NUMBER: US/10/414,692
; CURRENT FILING DATE: 2003-04-14
; PRIOR APPLICATION NUMBER: 60/372,650
; PRIOR FILING DATE: 2002-04-15
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 14
; LENGTH: 7554
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-414-692-14

Query Match      0.6%; Score 24; DB 15; Length 7554;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
Db 1507 GCTGCTGCTGCTGCTGCTGCTGCT 1484

RESULT 55
US-09-819-104A-4/c
; Sequence 4, Application US/09819104A
; Publication No. US20030027137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
; FILE REFERENCE: UMG-030
; CURRENT APPLICATION NUMBER: US/09/819,104A
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/193,138
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 8544
; TYPE: DNA

```

```

; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (160)..(7545)
US-09-819-104A-4

Query Match      0.6%; Score 24; DB 10; Length 8544;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
Db 1678 GCTGCTGCTGCTGCTGCTGCTGCT 1655

RESULT 56
US-10-174-014-11/c
; Sequence 11, Application US/10174014
; Publication No. US20040005292A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION
; FILE REFERENCE: PTS-0012
; CURRENT APPLICATION NUMBER: US/10/174,014
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 73
; SEQ ID NO 11
; LENGTH: 8561
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(7555)
US-10-174-014-11

Query Match      0.6%; Score 24; DB 15; Length 8561;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
Db 1508 GCTGCTGCTGCTGCTGCTGCTGCT 1485

RESULT 57
US-09-819-104A-1/c
; Sequence 1, Application US/09819104A
; Publication No. US20030027137A1
; GENERAL INFORMATION:
; APPLICANT: Chen, J. Don
; TITLE OF INVENTION: NOVEL NUCLEAR RECEPTOR COREPRESSOR MOLECULES
; FILE REFERENCE: UMG-030
; CURRENT APPLICATION NUMBER: US/09/819,104A
; CURRENT FILING DATE: 2001-03-27
; PRIOR APPLICATION NUMBER: 60/193,138
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8686
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (157)..(7677)
US-09-819-104A-1

Query Match      0.6%; Score 24; DB 10; Length 8686;
Best Local Similarity 100.0%; Pred. No. 0.28;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 180 GCTGCTGCTGCTGCTGCTGCTGCTGCT 203  
Db 1660 GCTGCTGCTGCTGCTGCTGCTGCTGCT 1637

## RESULT 58

US-10-174-014-4/c  
; Sequence 4, Application US/10174014  
; Publication No. US20040005292A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Susan M. Freier  
; APPLICANT: Kenneth W. Doble  
; TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION  
; FILE REFERENCE: PIS-0012  
; CURRENT APPLICATION NUMBER: US/10/174,014  
; CURRENT FILING DATE: 2002-06-17  
; NUMBER OF SEQ ID NOS: 73  
; SEQ ID NO 4  
; LENGTH: 8686  
; TYPE: DNA  
; ORGANISM: H. sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (157)...(7680)  
US-10-174-014-4

Query Match 0.6%; Score 24; DB 15; Length 8686;  
Best Local Similarity 100.0%; Pred. No. 0.28;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCTGCT 203  
Db 1660 GCTGCTGCTGCTGCTGCTGCTGCTGCT 1637

## RESULT 59

US-10-051-874-55/c  
; Sequence 55, Application US/10051874  
; Publication No. US2004000557A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Alsobrook II, John P  
; APPLICANT: Colman, Steven D  
; APPLICANT: Spytek, Kimberly A  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Vernet, Corine AM  
; APPLICANT: Li, Li  
; APPLICANT: Shency, Suresh G  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Guo, Xiaojia Sasha  
; APPLICANT: Edinger, Shlomit R  
; APPLICANT: MacDougall, John R  
; APPLICANT: Malyankar, Uriel M  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shinkets, Richard A  
; APPLICANT: Pena, Carol EA  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Zerhusen, Bryan D  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Miller, Charles E  
; APPLICANT: Lepley, Denise M  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Baumgartner, Jason C  
; APPLICANT: Herrman, John L  
; APPLICANT: Peyman, John A  
; APPLICANT: Gorman, Linda  
; APPLICANT: Mezes, Peter D  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Taupier Jr, Raymond J  
; APPLICANT: Gerlach, Valerie

; APPLICANT: Grosse, William M  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Rothenberg, Mark  
; APPLICANT: Stone, David J  
; APPLICANT: Burgess, Catherine E  
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF  
; FILE REFERENCE: 21402-245  
; CURRENT APPLICATION NUMBER: US/10/051,874  
; CURRENT FILING DATE: 2002-09-25  
; PRIOR APPLICATION NUMBER: 60/268,595.  
; PRIOR FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: 60/325,306  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 60/262,587  
; PRIOR FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: 60/272,409  
; PRIOR FILING DATE: 2001-02-28  
; PRIOR APPLICATION NUMBER: 60/262,454  
; PRIOR FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: 60/276,777  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/291,672  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/330,336  
; PRIOR FILING DATE: 2001-10-18  
; PRIOR APPLICATION NUMBER: 60/265,530  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/261,376  
; PRIOR FILING DATE: 2001-01-16  
; NUMBER OF SEQ ID NOS: 269  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 55  
; LENGTH: 14859  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-051-874-55

Query Match 0.6%; Score 24; DB 15; Length 14859;  
Best Local Similarity 100.0%; Pred. No. 0.29;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCTGCT 203  
Db 9082 GCTGCTGCTGCTGCTGCTGCTGCTGCT 9059

## RESULT 60

US-10-085-198-111/c  
; Sequence 111, Application US/10085198  
; Publication No. US20040009907A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook et al.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-279  
; CURRENT APPLICATION NUMBER: US/10/085,198  
; CURRENT FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 60/271,646  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 60/276,401  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/311,981  
; PRIOR FILING DATE: 2001-08-13  
; PRIOR APPLICATION NUMBER: 60/312,858  
; PRIOR FILING DATE: 2001-08-16  
; PRIOR APPLICATION NUMBER: 60/271,840  
; PRIOR FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: 60/277,324  
; PRIOR FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/286,096  
; PRIOR FILING DATE: 2001-04-21  
; PRIOR APPLICATION NUMBER: 60/299,695



```
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: 60/315,614
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/272,405
; PRIOR FILING DATE: 2001-02-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 653
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 111
; LENGTH: 15645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-198-111

Query Match      0.6%; Score 24; DB 15; Length 15645;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 GCTGCTGCTGCTGCTGTTGTGCT 203
Db 9082 GCTGCTGCTGCTGTTGTGCT 9059

RESULT 61
US-10-174-014-12/c
; Sequence 12, Application US/10174014
; Publication No. US20040008292A1
; GENERAL INFORMATION:
; APPLICANT: C. Frank Bennett
; APPLICANT: Susan M. Freier
; APPLICANT: Kenneth W. Dobie
; TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION
; FILE REFERENCE: PTS-0012
; CURRENT APPLICATION NUMBER: US/10/174,014
; CURRENT FILING DATE: 2002-06-17
; NUMBER OF SEQ ID NOS: 73
; SEQ ID NO 12
; LENGTH: 221000
; TYPE: DNA
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77967
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77968
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77969
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77970
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77971
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77972
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77973
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77974
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77975
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77976
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77977
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77978
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77979
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77980
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77981
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77982
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77983
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77984
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77985
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77986
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77987
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77988
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77989
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77990
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77991
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 77992
; OTHER INFORMATION: unknown
; FEATURE:
; NAME/KEY: unsure
```

LOCATION: 77993  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 77994  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 77995  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 77996  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 77997  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 77998  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 77999  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78000  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78001  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78002  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78003  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78004  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78005  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78006  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78007  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78008  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78009  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78010  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78011

OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78012  
OTHER INFORMATION: unknown  
FEATURE:  
Query Match 0.8%; Score 24; DB 15; Length 221000;  
Best Local Similarity 100.0%; Pred. No. 0.33;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203  
Db 137744 GCTGCTGCTGCTGCTGCTGCTGCT 137721  
RESULT 62  
US-09-893-238-24  
Sequence 24, Application US/09893238  
Patent No. US20020150973A1  
GENERAL INFORMATION:  
APPLICANT: Moore, K.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
FILE REFERENCE: 7853-237  
CURRENT APPLICATION NUMBER: US/09/893,238  
CURRENT FILING DATE: 2001-06-27  
PRIOR APPLICATION NUMBER: 09/245,041  
PRIOR FILING DATE: 1999-02-05  
PRIOR APPLICATION NUMBER: 60/093,630  
PRIOR FILING DATE: 1998-07-21  
PRIOR APPLICATION NUMBER: 60/104,978  
PRIOR FILING DATE: 1998-10-20  
NUMBER OF SEQ ID NOS: 129  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 24  
LENGTH: 29  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Primer  
US-09-893-238-24  
Query Match 0.5%; Score 23; DB 9; Length 29;  
Best Local Similarity 100.0%; Pred. No. 0.73;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 495 GCCAATAGATAATGAGACTTC 517  
Db 7 GCCAATAGATAATGAGACTTC 29  
RESULT 63  
US-10-260-238-1970  
Sequence 1970, Application US/10260238  
Publication No. US20040016025A1  
GENERAL INFORMATION:  
APPLICANT: Budworth, Paul R.  
APPLICANT: Moughamer, Todd G.  
APPLICANT: Briggs, Steven P.  
APPLICANT: Cooper, Bret  
APPLICANT: Glazebrook, Jane  
APPLICANT: Goff, Stephen A.  
APPLICANT: Katagiri, Fumiyaki  
APPLICANT: Kreps, Joel  
APPLICANT: Provart, Nicholas  
APPLICANT: Ricke, Darrell  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION  
FILE REFERENCE: 60111-NP  
CURRENT APPLICATION NUMBER: US/10/260,238  
CURRENT FILING DATE: 2002-09-26

```

, PRIOR APPLICATION NUMBER: US 60/325,448
, PRIOR FILING DATE: 2001-09-26
, PRIOR APPLICATION NUMBER: US 60/325,277
, PRIOR FILING DATE: 2001-09-26
, PRIOR APPLICATION NUMBER: US 60/370,620
, PRIOR FILING DATE: 2002-04-04
, NUMBER OF SEQ ID NOS: 6077
, SEQ ID NO 1970
, LENGTH: 104
, TYPE: DNA
, ORGANISM: Oryza sativa
, FEATURE:
, NAME/KEY: N region
, LOCATION: (1)..(11)
, OTHER INFORMATION: n = any nucleotide
, FEATURE:
, NAME/KEY: N region
, LOCATION: (79)..(79)
, OTHER INFORMATION: n = any nucleotide
, US-10-260-238-1970

```

```

Query Match      0.5%; Score 23; DB 15; Length 104;
Best Local Similarity 100.0%; Pred. No. 0.78;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      247  GCGCGCGCGCGCGCGCGCGT 269
          |||||
Db       80  GCGCGCGCGCGCGCGCGGT 102

```

	Query Match	0.5%	Score 23;	DB 9;	Length 221;
	Best Local Similarity	100.0%;	Pred. No. 0.8;		
	Matches 23; Conservative	0;	Mismatches	0;	Gaps 0;
Qy	2713	AAGGCTGCACCTGCATCAACC	2735		
Dd	118	AAGGCTGCACCTGCATCAACC	96		

```

; TITLE OF INVENTION:  DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE:  7853-237
; CURRENT APPLICATION NUMBER:  US/09/893,238
; CURRENT FILING DATE:  2001-06-27
; PRIOR APPLICATION NUMBER:  09/245,041
; PRIOR FILING DATE:  1999-02-05
; PRIOR APPLICATION NUMBER:  60/093,630
; PRIOR FILING DATE:  1998-07-21
; PRIOR APPLICATION NUMBER:  60/104,978
; PRIOR FILING DATE:  1998-10-20
; NUMBER OF SEQ ID NOS:  129
; SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 6
; LENGTH:  226
; TYPE:  DNA
; ORGANISM:  Mus musculus
US-09-893-238-6

```

Query Match	0.5%;	Score 23;	DB 9;	Length 226;
Best Local Similarity	100.0%;	Pred. No. 0.8;		
Matches 23;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2713	AAGGCTGCACCTGCATCAACC	2735	
Db	123	AAGGCTGCACCTGCATCAACC	101	

	Query Match	0.5%;	Score 23;	DB 14;	Length 500;
	Best local Similarity	100.0%;	Pred. No. 0.83;		
	Matches 23;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	210	GC GC GCT GTT TCT GCT GCT GTC TC	232		
Db	342	GC GC GCT GTT TCT GCT GCT GTC TC	364		

APPLICANT: Cooper, Bret  
APPLICANT: Glazebrook, Jane  
APPLICANT: Goff, Stephen A.  
APPLICANT: Katagiri, Fumiyaki  
APPLICANT: Kreps, Joel  
APPLICANT: Provart, Nicholas  
APPLICANT: Ricke, Darrell  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION  
FILE REFERENCE: 60111-NP  
CURRENT APPLICATION NUMBER: US/10/260,238  
CURRENT FILING DATE: 2002-09-26  
PRIOR APPLICATION NUMBER: US 60/325,448  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/325,277  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/370,620  
PRIOR FILING DATE: 2002-04-04  
NUMBER OF SEQ ID NOS: 6077  
SEQ ID NO 5694  
LENGTH: 510  
TYPE: DNA  
ORGANISM: Zea mays  
US-10-260-238-5694

Query Match 0.5%; Score 23; DB 15; Length 510;  
Best Local Similarity 100.0%; Pred.No. 0.63;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGGCGGCGGCGGCGGCGG 270  
DB 49 CCGCGGCGGCGGCGGCGGCGG 27

RESULT 68  
US-09-864-761-19977/c  
Sequence 19977, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aemica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: US 09/774,203  
PRIOR FILING DATE: 2001-01-29  
NUMBER OF SEQ ID NOS: 49117  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 19977  
LENGTH: 625  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC004827.1  
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.7  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.1  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4  
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.6  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.7  
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.5  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.8  
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.4  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5  
OTHER INFORMATION: EST HUMAN HIT: AW881551.1, EVALUE 6.00e-31  
OTHER INFORMATION: SWISSPROT HIT: QC4891, EVALUE 5.40e+00  
OTHER INFORMATION: NT HIT: AF286598.1, EVALUE 0.00e+00  
US-09-864-761-19977

Query Match 0.5%; Score 23; DB 9; Length 625;  
Best Local Similarity 100.0%; Pred.No. 0.84;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 GGCTGCTGCTGCTGCTGCTG 201  
DB 429 GGCTGCTGCTGCTGCTGCTG 407

RESULT 69  
US-10-424-599-110798/c  
Sequence 110798, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 110798  
LENGTH: 728  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
NAME/KEY: unsure  
LOCATION: (1)-(728)  
OTHER INFORMATION: unsure at all n locations  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_71060C.1  
US-10-424-599-110798

Query Match 0.5%; Score 23; DB 12; Length 728;  
Best Local Similarity 100.0%; Pred.No. 0.85;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 250 GCGGCGGCGGCGGCGGCGGTGC 272  
|||||  
Db 207 GCGGCGGCGGCGGCGGCGGTGC 185

## RESULT 70

US-10-425-114-20706  
; Sequence 20706, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kowalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 20706  
; LENGTH: 900  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3245-068-H12\_FLI  
US-10-425-114-20706

Query Match 0.5%; Score 23; DB 12; Length 900;  
Best Local Similarity 100.0%; Pred. No. 0.86;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 246 GCGCGCGCGCGCGCGCGCGG 268  
|||||  
Db 540 GCGCGCGCGCGCGCGCGCGG 562

## RESULT 71

US-10-260-238-2125  
; Sequence 2125, Application US/10260238  
; Publication No. US20040016025A1  
; GENERAL INFORMATION:  
; APPLICANT: Budworth, Paul R.  
; APPLICANT: Moughamer, Todd G.  
; APPLICANT: Briggs, Steven P.  
; APPLICANT: Cooper, Bret  
; APPLICANT: Glazebrook, Jane  
; APPLICANT: Goff, Stephen A.  
; APPLICANT: Katagiri, Fumiyaki  
; APPLICANT: Kreps, Joel  
; APPLICANT: Provart, Nicholas  
; APPLICANT: Ricke, Darrell  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION  
; FILE REFERENCE: 60111-NP  
; CURRENT APPLICATION NUMBER: US/10/260,238  
; CURRENT FILING DATE: 2002-09-26  
; PRIOR APPLICATION NUMBER: US 60/325,448  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: US 60/325,277  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: US 60/370,620  
; PRIOR FILING DATE: 2002-04-04  
; NUMBER OF SEQ ID NOS: 6077  
; SEQ ID NO 2125  
; LENGTH: 1320  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: N\_region

; LOCATION: (4)..(4)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: N\_region  
; LOCATION: (6)..(7)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: N\_region  
; LOCATION: (769)..(769)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: N\_region  
; LOCATION: (857)..(858)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: N\_region  
; LOCATION: (930)..(930)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: N\_region  
; LOCATION: (1042)..(1042)  
; OTHER INFORMATION: n = any nucleotide  
; FEATURE:  
; NAME/KEY: N\_region  
; LOCATION: (1044)..(1044)  
; OTHER INFORMATION: n = any nucleotide  
US-10-260-238-2125

Query Match 0.5%; Score 23; DB 15; Length 1320;  
Best Local Similarity 100.0%; Pred. No. 0.87;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 180 GCTGCTGCTGCTGCTGCTGTC 202  
|||||  
Db 1140 GCTGCTGCTGCTGCTGCTGTC 1162

## RESULT 72

US-09-864-761-3198/c  
; Sequence 3198, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeomica-x-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3198
; LENGTH: 1979
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004827.1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.7
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.5
; US-09-864-761-3198

Query Match 0.5%; Score 23; DB 9; Length 1979;
Best Local Similarity 100.0%; Pred.No. 0.89;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 GCGTCTGCTGCTGCTGCTGTTG 201
DB 733 GCGTCTGCTGCTGCTGTTG 711

RESULT 73
US-10-260-238-1864/c
; Sequence 1864, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Rumiya
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1864
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N region
; LOCATION: (489)..(489)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (501)..(501)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (549)..(549)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (552)..(552)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (554)..(554)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region

; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1814)..(1814)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1817)..(1817)
; OTHER INFORMATION: n = any nucleotide
; US-10-260-238-1864

Query Match 0.5%; Score 23; DB 15; Length 2000;
Best Local Similarity 100.0%; Pred.No. 0.89;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGCGCGCGGTTG 270
DB 1921 CCGCGCGCGCGCGCGGTTG 1899

RESULT 74
US-10-260-238-1920/c
; Sequence 1920, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Katagiri, Rumiya
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 1920
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N region
; LOCATION: (489)..(489)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (501)..(501)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (549)..(549)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (552)..(552)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (554)..(554)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region

```

```
; LOCATION: (559)...(559)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1235)...(1238)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1281)...(1281)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1300)...(1300)
; OTHER INFORMATION: n = any nucleotide
US-10-260-238-1920
```

```
Query Match 0.5%; Score 23; DB 15; Length 2000;
Best Local Similarity 100.0%; Pred.No. 0.89;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 250 GCGGCGGCGGCGGCGGTGTC 272
      |||||||
Db 307 GCGGCGGCGGCGGCGGTGTC 285
```

```
RESULT 75
US-10-108-260A-790/c
; Sequence 790, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 790
; LENGTH: 2264
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-790
```

```
Query Match 0.5%; Score 23; DB 15; Length 2264;
Best Local Similarity 100.0%; Pred.No. 0.89;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 210 GCGGCTGTGTGCTGCTGCTGC 232
      |||||||
Db 564 GCGGCTGTGTGCTGCTGCTGC 542
```

```
Search completed: March 1, 2004, 21:14:28
Job time : 1419 secs
```

**This Page Blank (uspto)**





## ALIGNMENTS

```

RESULT 1
US-09-245-041-14
; Sequence 14, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 8589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-14

```

Query Match	88.0%;	Score 3774;	DB 3;	Length 8589;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 3974;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;
QY	313	TGTTCAACGGCGGTGCGTCTCAACCTTGGCACGGCCAGTGCCTCTGCCCGCGCGTGG	372	
DB	310	TGTTCAACGGCGGTGCGTCTCAACCTTGGCACGGCCAGTGCCTCTGCCCGCGCGTGG	369	
QY	373	GTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAACTGATCTTCTGGGTTT	432	
DB	370	GTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAACTGATCTTCTGGGTTT	429	
QY	433	GTGACGATCGACTCTGAAATTTATAATACAAACGAAGTGCACGTGGCTCAITTAAGGA	492	
DB	430	GTGACGATCGACTCTGAAATTTATAATACAAACGAAGTGCACGTGGCTCAITTAAGGA	489	
QY	493	CAGCCAAATAGATAATAGAGACTTGGTTTCAATCATTTTCCCTACAGAGTGTAGTGGAC	552	
DB	490	CAGCCAAATAGATAATAGAGACTTGGTTTCAATCATTTTCCCTACAGAGTGTAGTGGAC	549	
QY	553	CATTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGC	612	
DB	550	CATTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGC	609	
QY	613	CTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT	672	
DB	610	CTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGT	669	
QY	673	TATGCCCTTGCTGCATTTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACT	732	
DB	670	TATGCCCTTGCTGCATTTTTTTTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACT	729	
QY	733	TACAGTTTTCATATGTTGCCAATAACTGCTCAGGCCGAGGAGTGTAGATCAGTAAT	792	
DB	730	TACAGTTTTCATATGTTGCCAATAACTGCTCAGGCCGAGGAGTGTAGATCAGTAAT	789	
QY	793	AGCAGCGAAACTGTGGAATGTAATGTTCTGAAAACTGAAAAAGTGAAGCATGTGCAATT	852	
DB	790	AGCAGCGAATCTGTGGAATGTAATGTTCTGAAAACTGAAAAAGTGAAGCATGTGCAATT	849	
QY	853	CCTCCTGTACAGCAACTGTGGTTTTTTCCTCATCGAGGCAATCTGCAATTCAGTGTATGTC	912	
DB	850	CCTCCTGTACAGCAACTGTGGTTTTTTCCTCATCGAGGCAATCTGCAATTCAGTGTATGTC	909	





1933 TCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTGTTTGTAGCAGCAGCAGCCTGCTATT 1992  
1930 TCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTGTTTGTAGCAGCAGCAGCCTGCTATT 1989  
1993 CGGTGTGTGTGGAACACAGAGTGTCTCTCAGTGTATCTCGTGGGCGGTGGCAACTGATGAA 2052  
1990 CGGTGTGTGTGGAACACAGAGTGTCTCTCAGTGTATCTCGTGGGCGGTGGCAACTGATGAA 2049  
2053 CAAGAAGAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTCTTGACCAATGACAGATGT 2112  
2050 CAAGAAGAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTCTTGACCAATGACAGATGT 2109  
2113 GACCAGACACAGATGTTACAGCTGTACAGCTGTACAGCAACCAATGACTGCCACTGGTGAAT 2172  
2110 GACCAGACACAGATGTTTACAGCTGTACAGCTGTACAGCAACCAATGACTGCCACTGGTGAAT 2169  
2173 GACCATTGTGTCCCGAGAAACCAAGCTGTCTGAGAGGCCAGATCTCCATTTTGTAGTAT 2232  
2170 GACCATTGTGTCCCGAGAAACCAAGCTGTCTGAGAGGCCAGATCTCCATTTTGTAGTAT 2229  
2233 GAGAAATTTGCCCGAGAAACCAAGCTGTCTGAGAGGCCAGATCTCCATTTTGTAGTAT 2292  
2230 GAGAAATTTGCCCGAGAAACCAAGCTGTCTGAGAGGCCAGATCTCCATTTTGTAGTAT 2289  
2293 TGTGCTTGGACACAGAACTGCGAGTGGAGGCCCGGAATCAGGAGTGCATTTGCCCTGCC 2352  
2290 TGTGCTTGGACACAGAACTGCGAGTGGAGGCCCGGAATCAGGAGTGCATTTGCCCTGCC 2349  
2353 GAAATATCTGTGCAATTTGGCTGCAATTTGGTGGAACTCACTGTTTGAATTTACTACT 2412  
2350 GAAATATCTGTGCAATTTGGCTGCAATTTGGTGGAACTCACTGTTTGAATTTACTACT 2409  
2413 GCCAAGGAGAAATTAACAAATGCTTAAATGTTCTGTAGGAACCAATGCGCTTTGGCT 2472  
2410 GCCAAGGAGAAATTAACAAATGCTTAAATGTTCTGTAGGAACCAATGCGCTTTGGCT 2469  
2473 TCTCTTAAACCCAGAGAAAGTGAATTTGCTCTTAAAGCAGCTGCGGAATTAAGCAGTCA 2532  
2470 TCTCTTAAACCCAGAGAAAGTGAATTTGCTCTTAAAGCAGCTGCGGAATTAAGCAGTCA 2529  
2533 TCTCAGAGCATGTCAGCTCACTTAACCCATGCGGTGCGCTTCGGAAGATCAATGTG 2592  
2530 TCTCAGAGCATGTCAGCTCACTTAACCCATGCGGTGCGCTTCGGAAGATCAATGTG 2589  
2593 TCCTACTGCTGTGGAGATATGTCCTTCAATTAACAAATGTTTACTACAGTGGATCCG 2652  
2590 TCCTACTGCTGTGGAGATATGTCCTTCAATTAACAAATGTTTACTACAGTGGATCCG 2649  
2653 TCTGAGCCAGTGTGATGCTGGAATCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 2712  
2650 TCTGAGCCAGTGTGATGCTGGAATCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTG 2709  
2713 AAGGCTCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCCTGCAAAACAC 2772  
2710 AAGGCTCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCCTGCAAAACAC 2769  
2773 AGTGCTAAGAGTGTCCGACACCACTGTGCTTGAAGCAGCATGTGGAGATGACACAGC 2832  
2770 AGTGCTAAGAGTGTCCGACACCACTGTGCTTGAAGCAGCATGTGGAGATGACACAGC 2829  
2833 GGCAGCTGTGAGTGTGATGCTGTGAGCAACATGAAGCAGTGTGGAATCCCAATGCTAT 2892  
2830 GGCAGCTGTGAGTGTGATGCTGTGAGCAACATGAAGCAGTGTGGAATCCCAATGCTAT 2889  
2893 GTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACTGCCGCCCT 2952  
2890 GTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACTGCCGCCCT 2949  
2953 GAAATTTGTTCAGGCTACTGATGCTGATGCTGCTTGGAGCAACCAAGGCTGTGGCTGG 3012  
2950 GAAATTTGTTCAGGCTACTGATGCTGATGCTGCTTGGAGCAACCAAGGCTGTGGCTGG 3009  
3013 TGTACTGATCCCGACCAATACTGGCAAGGAAATGCAATAGAGGGTTCCTATAAAGGACCA 3072

3010 TGTACTGATCCAGCAATACTGGCAAGGAAATGCAATAGAGGGTTCCTATAAAGGACCA 3069  
3073 GTGAAGATGCTTTCGCAAGCCCTTACAGAAATTTCTATCCACAGCCCTGCTCAATTC 3132  
3070 GTGAAGATGCTTTCGCAAGCCCTTACAGAAATTTCTATCCACAGCCCTGCTCAATTC 3129  
3133 AGCATGTGTCTAGAGCAGCAGATACAACTGGTCTTTTCACTCTGCTCCAGCTTGCCAA 3192  
3130 AGCATGTGTCTAGAGCAGCAGATACAACTGGTCTTTTCACTCTGCTCCAGCTTGCCAA 3189  
3193 TGCACCGGCCACAGTAAATGCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGACC 3252  
3190 TGCACCGGCCACAGTAAATGCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGACC 3249  
3253 ACAGCAAGCATCTCGAGACCTGCATATCTGCTTCTAGGTTAGTCCACCAATGAGGG 3312  
3250 ACAGCAAGCATCTCGAGACCTGCATATCTGCTTCTAGGTTAGTCCACCAATGAGGG 3309  
3313 AAATGTGACCCATGCAAGTGGCAACGCGCTCTCTGTGCAACCAACCAACGCGCAAG 3372  
3310 AAATGTGACCCATGCAAGTGGCAACGCGCTCTCTGTGCAACCAACCAACGCGCAAG 3369  
3373 TGTCTTGCACCAACCAAGGCGCTCAAGGGGACGAGTCCAGCTATGTGAGTGAATAT 3432  
3370 TGTCTTGCACCAACCAAGGCGCTCAAGGGGACGAGTCCAGCTATGTGAGTGAATAT 3429  
3433 CGATACCAAGGAAACCCCTCTCAGAGGAAATCTTATATCTCTTCTTATTCAGTATCAG 3492  
3430 CGATACCAAGGAAACCCCTCTCAGAGGAAATCTTATATCTCTTCTTATTCAGTATCAG 3489  
3493 TTCACCTTTAGTCTATCCAGGAGATGATGCTTATACAGACTATCAATTTTGGCT 3552  
3490 TTCACCTTTAGTCTATCCAGGAGATGATGCTTATACAGACTATCAATTTTGGCT 3549  
3553 ACTCTGACGAAACCAACAGGATTTGGACATGTTTCATCAATGCTCTCCAAAGATTTCAAC 3612  
3550 ACTCTGACGAAACCAACAGGATTTGGACATGTTTCATCAATGCTCTCCAAAGATTTCAAC 3609  
3613 CTCACATCAGCTGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCT 3672  
3610 CTCACATCAGCTGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCT 3669  
3673 GTTGTCTTCAAAACCAACATTAAGGATGAAGATAGTTTCTTAATCAGAGTTTGTAT 3732  
3670 GTTGTCTTCAAAACCAACATTAAGGATGAAGATAGTTTCTTAATCAGAGTTTGTAT 3729  
3733 TTTGCAACCAACCAATATCACTTTCTTTTGTATGTGAGTAAATTTACCTGGCCCATC 3792  
3730 TTTGCAACCAACCAATATCACTTTCTTTTGTATGTGAGTAAATTTACCTGGCCCATC 3789  
3793 AAAATTCAG 3801  
3790 AAAATTCAG 3798

RESULT 3

US-09-245-041-18  
; Sequence 18, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 18

; LENGTH: 2625

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-245-041-18

## Query Match

Best Local Similarity 42.8%; Score 1837; DB 3; Length 2625;

Matches 2037; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	313	TGTGTCAACGGCGGTGCTGCAACCGTCGACCGCGGAGTGGTGTGCGCGCGCGGTGG	372
DB	310	TGTGTCAACGGCGGTGCTGCAACCGTCGACCGCGGAGTGGTGTGCGCGCGGTGG	369
QY	373	GTGGCGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAACTGGGATCTTCTGGGTTT	432
DB	370	GTGGCGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAACTGGGATCTTCTGGGTTT	429
QY	433	GTGACAGATGGACCTGGGAAATTAATAACAAACGAAAGTGCAGTGGCTCAITGAAGGA	492
DB	430	GTGACAGATGGACCTGGGAAATTAATAACAAACGAAAGTGCAGTGGCTCAITGAAGGA	489
QY	493	CAGCCAAATAGAAATAGAGACTTGGTTCATCAATTTTGTACAGAGTGTAGTGGAC	552
DB	490	CAGCCAAATAGAAATAGAGACTTGGTTCATCAATTTTGTACAGAGTGTAGTGGAC	549
QY	553	CATTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTGCAATTTAGTGGC	612
DB	550	CATTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTGCAATTTAGTGGC	609
QY	613	CTCATGTTCTGAGAGATGGCAATCAGACTGCTCCCTGAGGTGTGGACATCAGGT	672
DB	610	CTCATGTTCTGAGAGATGGCAATCAGACTGCTCCCTGAGGTGTGGACATCAGGT	669
QY	673	TATGCTTGTGCTGCAATTTTATGATGCTGCTTATTAATTTGACTGGATTTAATTAAT	732
DB	670	TATGCTTGTGCTGCAATTTTATGATGCTGCTTATTAATTTGACTGGATTTAATTAAT	729
QY	733	TACAGTTTGTATGTGTCGCAATTAATGCTGACCGCGAGAGAGTGTAAAGATCAGTAAT	792
DB	730	TACAGTTTGTATGTGTCGCAATTAATGCTGACCGCGAGAGAGTGTAAAGATCAGTAAT	789
QY	793	AGCAGCGAACTGTTGAATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACAT	852
DB	790	AGCAGCGAATCTGTTGAATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACAT	849
QY	853	CCTCACTGTACAGACAACTGTGGTTTCTCATCGAGGCACTGTGCAATTTCAAGTGTGTC	912
DB	850	CCTCACTGTACAGACAACTGTGGTTTCTCATCGAGGCACTGTGCAATTTCAAGTGTGTC	909
QY	913	AGAGATGCTCTGCTTCTCAGACTGCGAGGTCCTGGATGTTGAGTTCCTGTACAGCT	972
DB	910	AGAGATGCTCTGCTTCTCAGACTGCGAGGTCCTGGATGTTGAGTTCCTGTACAGCT	969
QY	973	AACAGTCAATTTGGACTCGAGAGGAATATTCTAACTTAAAGCTTCCCGAGAGATCTCAT	1032
DB	970	AACAGTCAATTTGGACTCGAGAGGAATATTCTAACTTAAAGCTTCCCGAGAGATCTCAT	1029
QY	1033	AAAGCTGTGTCATGGAACAAATTTATGTTGGTGTGTTGAGGATATATGTTCAACCACTCA	1092
DB	1030	AAAGCTGTGTCATGGAACAAATTTATGTTGGTGTGTTGAGGATATATGTTCAACCACTCA	1089
QY	1093	GATTATAACATGGTTCAGCTGTGCTTCTAGGGAGTGGCTTCCACTAAACCGT	1152
DB	1090	GATTATAACATGGTTCAGCTGTGCTTCTAGGGAGTGGCTTCCACTAAACCGT	1149
QY	1153	TCTGTGAACAAATGTGTTGTAGTATGTTGATCTTCTTGGCATATATCAAGGATAAAAT	1212
DB	1150	TCTGTGAACAAATGTGTTGTAGTATGTTGATCTTCTTGGCATATATCAAGGATAAAAT	1209
QY	1213	TACATGTATGAGGAAATTTGATCCAACTGGGAAATGTGACCAATGATGTTGAGATTTT	1272

DB	1210	TACATGTATGAGGAAAAATTTGATTCAACTGGGAATGTGACCAATGAGTTGAGATTTT	1269
QY	1273	CACATTCAATAGATCATGGGTGTTGTTGACCCCTAAGGCAAGGAGCAGTATGCAATG	1332
DB	1270	CACATTCAATAGATCATGGGTGTTGTTGACCCCTAAGGCAAGGAGCAGTATGCAATG	1329
QY	1333	GTGGGCACTCTGCAACACATTTGTTACCTGAAGAATGGCGGAGTGTGATGCTGTCATC	1392
DB	1330	GTGGGCACTCTGCAACACATTTGTTACCTGAAGAATGGCGGAGTGTGATGCTGTCATC	1389
QY	1393	TTTGGTCACTGCCCTCTCTATGGAATATAAGGAATGTGCAAGAAATGATTTGGATAG	1452
DB	1390	TTTGGTCACTGCCCTCTCTATGGAATATAAGGAATGTGCAAGAAATGATTTGGATAG	1449
QY	1453	AACACATGGAGTATATTACACACCCAGGGTGCCTTTGTGCAAGGGGTTACGGCCATAGC	1512
DB	1450	AACACATGGAGTATATTACACACCCAGGGTGCCTTTGTGCAAGGGGTTACGGCCATAGC	1509
QY	1513	AGTGTTTAGCAACATAGGACAGGGCCCTATACGTTTATGTTGGCTACAGGCTTTTCACT	1572
DB	1510	AGTGTTTAGCAACATAGGACAGGGCCCTATACGTTTATGTTGGCTACAGGCTTTTCACT	1569
QY	1573	GCCAAATAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGATACCCAGATGCG	1632
DB	1570	GCCAAATAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGATACCCAGATGCG	1629
QY	1633	ACCATTCTTTAAGGACAGCGGATTTTTCGGTTACTTTGCAACACAGCTGTGATAGTGGGA	1692
DB	1630	ACCATTCTTTAAGGACAGCGGATTTTTCGGTTACTTTGCAACACAGCTGTGATAGTGGGA	1689
QY	1693	ACCATGCTGCTGTTGGGGAAACACACACANTGACATCTATGAGGCATGCGGCCAAA	1752
DB	1690	ACCATGCTGCTGTTGGGGAAACACACACANTGACATCTATGAGGCATGCGGCCAAA	1749
QY	1753	TGCTTCTCTTCAGATTTTATGCGCTTATGACATTTGCTGACCGCTGTGCTGCTTCCC	1812
DB	1750	TGCTTCTCTTCAGATTTTATGCGCTTATGACATTTGCTGACCGCTGTGCTGCTTCCC	1809
QY	1813	AGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTTAGAGTCTTTACACAAACG	1872
DB	1810	AGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTTAGAGTCTTTACACAAACG	1869
QY	1873	ACCATGATGCTGCTCGGTGTTTCAATAGTCTCTCCTCAGCGACATCTTGATTTACC	1932
DB	1870	ACCATGATGCTGCTCGGTGTTTCAATAGTCTCTCCTCAGCGACATCTTGATTTACC	1929
QY	1933	TCGGAACAGTGTGATGCGCATCGGAGTGAAGCGCTTTTGTAGCAGGACCTGGTATT	1992
DB	1930	TCGGAACAGTGTGATGCGCATCGGAGTGAAGCGCTTTTGTAGCAGGACCTGGTATT	1989
QY	1993	CGGTGTGTGGAACACAGGGTCTCAGTGTATCTCGTGGCGCTGCGCACTGATGAA	2052
DB	1990	CGGTGTGTGGAACACAGGGTCTCAGTGTATCTCGTGGCGCTGCGCACTGATGAA	2049
QY	2053	CAAGAAGAAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTTTTGACCATGACAGATGT	2112
DB	2050	CAAGAAGAAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTTTTGACCATGACAGATGT	2109
QY	2113	GACGAGCAACAGATTTGTTACGCTGTACAGGCAACACCAATGACTGCGCTGCGCAAT	2172
DB	2110	GACGAGCAACAGATTTGTTACGCTGTACAGGCAACACCAATGACTGCGCTGCGCAAT	2169
QY	2173	GACCAATGTGTCGCCAGGAAACACAGCTGCTCAGAAGGCCAGATCTCCATTTTAGGTAT	2232
DB	2170	GACCAATGTGTCGCCAGGAAACACAGCTGCTCAGAAGGCCAGATCTCCATTTTAGGTAT	2229
QY	2233	GAGAAATGCGCCCAAGGATAACCCCATGACTACTGTATTAACAAGAACACAGCTGAGGAGC	2292
DB	2230	GAGAAATGCGCCCAAGGATAACCCCATGACTACTGTATTAACAAGAACACAGCTGAGGAGC	2289
QY	2293	TGTGCTCTGGACCAAGAACTGCGAGTGGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCC	2352
DB	2290	TGTGCTCTGGACCAAGAACTGCGAGTGGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCC	2349

QY	2353	G	2353
DB	2350	G	2350
RESULT 4			
US-09-245-041-12			
; Sequence 12, Application US/09245041			
; Patent No. 6274339			
; GENERAL INFORMATION:			
; APPLICANT: Moore, K.			
; APPLICANT: Nagle, D.			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT			
; FILE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY			
; FILE REFERENCE: 7853-136			
; CURRENT APPLICATION NUMBER: US/09/245,041			
; CURRENT FILING DATE: 1999-02-05			
; EARLIER APPLICATION NUMBER: 60/093,630			
; EARLIER FILING DATE: 1998-07-21			
; EARLIER APPLICATION NUMBER: 60/104,978			
; EARLIER FILING DATE: 1998-10-20			
; NUMBER OF SEQ ID NOS: 131			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 12			
; LENGTH: 6370			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-245-041-12			
Query Match 35.2%; Score 1510; DB 3; Length 6370;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1510; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	2781	GCAGTGC	CGGACACCATGTGCTTGAGCAGCATGTGGAGATTGCACACGCGCAGCTC 2840
DB	548	GCAGTGC	CGGACACCATGTGCTTGAGCAGCATGTGGAGATTGCACACGCGCAGCTC 607
QY	2841	TCAGTGC	ATGTGCTGCACACATCAAGCAGTGTGGACTCCATGCTATGTGGCCTC 2900
DB	608	TCAGTGC	ATGTGCTGCACACATCAAGCAGTGTGGACTCCATGCTATGTGGCCTC 667
QY	2901	CTTCCCT	TTTGGCCAGTGTATGGAAATGATACGATGACGACCTGCCCCCTCGAAATTTG 2960
DB	668	CTTCCCT	TTTGGCCAGTGTATGGAAATGATACGATGACGACCTGCCCCCTCGAAATTTG 727
QY	2961	TTGAGCT	ACTCTACTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGTGCTACTGA 3020
DB	728	TTGAGCT	ACTCTACTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGTGCTACTGA 787
QY	3021	TCCCAGC	AACTACTGGCAAGGGAATGCATAGAGGTTTCTATAAGGACCAAGTCAAGAT 3080
DB	788	TCCCAGC	AACTACTGGCAAGGGAATGCATAGAGGTTTCTATAAGGACCAAGTCAAGAT 847
QY	3081	GCCTTGC	CAAGCCCTTACAGGAAATTTCTATCCAAGCCCTGTGCTCAATTCAGCATGTG 3140
DB	848	GCCTTGC	CAAGCCCTTACAGGAAATTTCTATCCAAGCCCTGTGCTCAATTCAGCATGTG 907
QY	3141	TCTAGAG	CACAGATCAACTGTCTTTCATTCACGTGTCAGCTTGCCAAATGCAACGG 3200
DB	908	TCTAGAG	CACAGATCAACTGTCTTTCATTCACGTGTCAGCTTGCCAAATGCAACGG 967
QY	3201	CCACAGT	AAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACACAGGCAA 3260
DB	968	CCACAGT	AAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACACAGGCAA 1027
QY	3261	GCACCTG	CAGACCTGATCTGGCTTCTACGGTGTATCCACCAATGGAGGAATGTCA 3320
DB	1028	GCACCTG	CAGACCTGATCTGGCTTCTACGGTGTATCCACCAATGGAGGAATGTCA 1087
QY	3321	GCCATCA	AGTGGCAACGCGTCTCTGTGCAACACCAACAGGCAAGTGTCTTG 3380
DB	1088	GCCATCA	AGTGGCAACGCGTCTCTGTGCAACACCAACAGGCAAGTGTCTTG 1147

QY	3381	CACCACCA	AGGGCGTCAAGGGGAGCAGTGCAGCTATGTGAGGTAGAAAAATCGATACCA 3440
DB	1148	CACCACCA	AGGGCGTCAAGGGGAGCAGTGCAGCTATGTGAGGTAGAAAAATCGATACCA 1207
QY	3441	AGGAAAC	CCCTCTCAGAGGAACATGTTATTATCTCTTCTATTGACTATCAGTTACCTT 3500
DB	1208	AGGAAAC	CCCTCTCAGAGGAACATGTTATTATCTCTTCTATTGACTATCAGTTACCTT 1267
QY	3501	TAGTCTAT	CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCTGGA 3560
DB	1268	TAGTCTAT	CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCTGGA 1327
QY	3561	CGAACAAC	AAACAGGATTTGACATGTTCAATGCTCCAAGATTTTCAACCTCAACAT 3620
DB	1328	CGAACAAC	AAACAGGATTTGACATGTTCAATGCTCCAAGATTTTCAACCTCAACAT 1387
QY	3621	CACCTGG	CGTCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTTGTTTC 3680
DB	1388	CACCTGG	CGTCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTTGTTTC 1447
QY	3681	AAAAACA	CAATTAAGAGTACAAAGATAGTTTCTTAATGAGAGTTTGTATTCGCA 3740
DB	1448	AAAAACA	CAATTAAGAGTACAAAGATAGTTTCTTAATGAGAGTTTGTATTCGCA 1507
QY	3741	CCACCAAA	TATCATCTTCTTGTATTGTCAATAATTTTCACTGGCCCATCAAAATTC 3800
DB	1508	CCACCAAA	TATCATCTTCTTGTATTGTCAATAATTTTCACTGGCCCATCAAAATTC 1567
QY	3801	GATTGCT	TCTCTCAGCAGCAAGCAATTTTATGACCTGGTACATGTTCTTGTGATCTT 3860
DB	1568	GATTGCT	TCTCTCAGCAGCAAGCAATTTTATGACCTGGTACATGTTCTTGTGATCTT 1627
QY	3861	CAGTGTTC	CTCTCTGCTGCTGCTGCTGTTTGGAGATCAAAACAAAGTTG 3920
DB	1628	CAGTGTTC	CTCTCTGCTGCTGCTGCTGTTTGGAGATCAAAACAAAGTTG 1687
QY	3921	TTGGGCT	CTCCAGAGTACCAACTTCTTCAGAGATGCAACAGATGGCCAGCCGCTCC 3980
DB	1688	TTGGGCT	CTCCAGAGTACCAACTTCTTCAGAGATGCAACAGATGGCCAGCCGCTCC 1747
QY	3981	CTTTGCC	CTGTAAATGTGCTTGGAAACAGATGAGGACCTCTCTGATCTTATTTGGGG 4040
DB	1748	CTTTGCC	CTGTAAATGTGCTTGGAAACAGATGAGGACCTCTCTGATCTTATTTGGGG 1807
QY	4041	GAGTATA	AAAGACTGTTCCCAAAACCCATTGCACTGGAGCCGTTTGGCAACAAAGCCG 4100
DB	1808	GAGTATA	AAAGACTGTTCCCAAAACCCATTGCACTGGAGCCGTTTGGCAACAAAGCCG 1867
QY	4101	TGTCCTC	CTCTGTGTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCTCTCTGGGCA 4160
DB	1868	TGTCCTC	CTCTGTGTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCTCTCTGGGCA 1927
QY	4161	GTGAGGT	CTTGTGTCAGCCCTCTGTGGACATTTCTCAGCAGATGCCGATAGTGA 4220
DB	1928	GTGAGGT	CTTGTGTCAGCCCTCTGTGGACATTTCTCAGCAGATGCCGATAGTGA 1987
QY	4221	CAAGGAG	AAAGTTCAGGACCCGTGAGAAACCGGAGAGCAGCCCTCTGCAAGCTGGGAC 4280
DB	1988	CAAGGAG	AAAGTTCAGGACCCGTGAGAAACCGGAGAGCAGCCCTCTGCAAGCTGGGAC 2047
QY	4281	CTGCAT	CTGA 4290
DB	2048	CTGCAT	CTGA 2057

RESULT 5  
US-09-245-041-1  
; Sequence 1, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.







```
RESULT 8
US-09-614-124B-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-966

Query Match      5.7%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.3e-112;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGACGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGC 2574
DB 246 CTGCGAATAATGACGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGC 187

QY 2575 CTTCCGAAGATCAATGTGTCCTACTGCTGGGAAGATATGTCCCATTTACAAATAGT 2634
DB 186 CTTCCGAAGATCAATGTGTCCTACTGCTGGGAAGATATGTCCCATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCCGCTGAGCCAGTCAGTGTGGAAATTTTATCAGAA 2694
DB 126 TTACTACAGTGGATGCCGCTGAGCCAGTCAGTGTGGAAATTTTATCAGAA 67

QY 2695 CCAGTACTCGGGGACTGAAGGCTCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
DB 66 CCAGTACTCGGGGACTGAAGGCTCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
DB 6 GAAAGG 1

RESULT 9
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0

Query Match      5.7%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.3e-112;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGACGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGC 2574
DB 246 CTGCGAATAATGACGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGC 187

QY 2575 CTTCCGAAGATCAATGTGTCCTACTGCTGGGAAGATATGTCCCATTTACAAATAGT 2634
DB 186 CTTCCGAAGATCAATGTGTCCTACTGCTGGGAAGATATGTCCCATTTACAAATAGT 127

QY 2635 TTACTACAGTGGATGCCGCTGAGCCAGTCAGTGTGGAAATTTTATCAGAA 2694
DB 126 TTACTACAGTGGATGCCGCTGAGCCAGTCAGTGTGGAAATTTTATCAGAA 67

QY 2695 CCAGTACTCGGGGACTGAAGGCTCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2754
DB 66 CCAGTACTCGGGGACTGAAGGCTCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2755 GAAAGG 2760
DB 6 GAAAGG 1

RESULT 9
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0

Query Match      1.3%; Score 55; DB 3; Length 90050;
Best Local Similarity 100.0%; Pred. No. 5.5e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3130 TCCAGCATGTCTTAGAGGACAGCATACAACCTGGTCTTTCATTCACTGTCCAG 3184
DB 31840 TCCAGCATGTCTTAGAGGACAGCATACAACCTGGTCTTTCATTCACTGTCCAG 31894

RESULT 11
US-09-245-041-10
; Sequence 10, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-5

; Sequence 5, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-5
```

FILE REFERENCE: 7853-136  
 CURRENT APPLICATION NUMBER: US/09/245,041  
 CURRENT FILING DATE: 1999-02-05  
 EARLIER APPLICATION NUMBER: 60/093,630  
 EARLIER FILING DATE: 1998-07-21  
 EARLIER APPLICATION NUMBER: 60/104,978  
 EARLIER FILING DATE: 1998-10-20  
 NUMBER OF SEQ ID NOS: 131  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 10  
 LENGTH: 1051  
 TYPE: DNA  
 ORGANISM: Mus musculus  
 US-09-245-041-10

Query Match 1.2%; Score 53; DB 3; Length 1051;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-16;  
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 GCGTGGTGGCGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAACTGG 419  
 Db 571 GCGTGGTGGCGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAACTGG 623

## RESULT 12

US-09-245-041-8  
 Sequence 8, Application US/09245041  
 Patent No. 6274339  
 GENERAL INFORMATION:  
 APPLICANT: Moore, K.  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
 OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
 FILE REFERENCE: 7853-136  
 CURRENT APPLICATION NUMBER: US/09/245,041  
 CURRENT FILING DATE: 1999-02-05  
 EARLIER APPLICATION NUMBER: 60/093,630  
 EARLIER FILING DATE: 1998-07-21  
 EARLIER APPLICATION NUMBER: 60/104,978  
 EARLIER FILING DATE: 1998-10-20  
 NUMBER OF SEQ ID NOS: 131  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 8  
 LENGTH: 2419  
 TYPE: DNA  
 ORGANISM: Mus musculus  
 US-09-245-041-8

Query Match 1.1%; Score 47; DB 3; Length 2419;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-13;  
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1624 CAGATGTGACCAATCTTAAGGACAGCGGATTTTCCGTTACTTGCA 1670  
 Db 1890 CAGATGTGACCAATCTTAAGGACAGCGGATTTTCCGTTACTTGCA 1936

## RESULT 13

US-09-245-041-4  
 Sequence 4, Application US/09245041  
 Patent No. 6274339  
 GENERAL INFORMATION:  
 APPLICANT: Moore, K.  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
 OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
 FILE REFERENCE: 7853-136  
 CURRENT APPLICATION NUMBER: US/09/245,041  
 CURRENT FILING DATE: 1999-02-05  
 EARLIER APPLICATION NUMBER: 60/093,630  
 EARLIER FILING DATE: 1998-07-21  
 EARLIER APPLICATION NUMBER: 60/104,978  
 EARLIER FILING DATE: 1998-10-20

NUMBER OF SEQ ID NOS: 131  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 4  
 LENGTH: 5973  
 TYPE: DNA  
 ORGANISM: Mus musculus  
 US-09-245-041-4

Query Match 0.8%; Score 35; DB 3; Length 5973;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-07;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 547 TGGGACCATTATATGTTTATGATGGGACTCAAT 581  
 Db 4379 TGGGACCATTATATGTTTATGATGGGACTCAAT 4413

## RESULT 14

US-09-245-041-3/c  
 Sequence 3, Application US/09245041  
 Patent No. 6274339  
 GENERAL INFORMATION:  
 APPLICANT: Moore, K.  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
 OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
 FILE REFERENCE: 7853-136  
 CURRENT APPLICATION NUMBER: US/09/245,041  
 CURRENT FILING DATE: 1999-02-05  
 EARLIER APPLICATION NUMBER: 60/093,630  
 EARLIER FILING DATE: 1998-07-21  
 EARLIER APPLICATION NUMBER: 60/104,978  
 EARLIER FILING DATE: 1998-10-20  
 NUMBER OF SEQ ID NOS: 131  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 3  
 LENGTH: 17056  
 TYPE: DNA  
 ORGANISM: Mus musculus  
 US-09-245-041-3

Query Match 0.7%; Score 32; DB 3; Length 17056;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-05;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 844 TGTGACATTCCTCACTGTACAGCAACTGTGG 875  
 Db 3978 TGTGACATTCCTCACTGTACAGCAACTGTGG 3947

## RESULT 15

US-09-245-041-23/c  
 Sequence 23, Application US/09245041  
 Patent No. 6274339  
 GENERAL INFORMATION:  
 APPLICANT: Moore, K.  
 TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
 OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
 FILE REFERENCE: 7853-136  
 CURRENT APPLICATION NUMBER: US/09/245,041  
 CURRENT FILING DATE: 1999-02-05  
 EARLIER APPLICATION NUMBER: 60/093,630  
 EARLIER FILING DATE: 1998-07-21  
 EARLIER APPLICATION NUMBER: 60/104,978  
 EARLIER FILING DATE: 1998-10-20  
 NUMBER OF SEQ ID NOS: 131  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 23  
 LENGTH: 29  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 US-09-245-041-23

Query Match 0.7%; Score 29; DB 3; Length 29;  
Best Local Similarity 100.0%; Pred. No. 0.00032;  
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 552 CCATTATATGTTTATGATGGGACTCAA 580  
DB 29 CCATTATATGTTTATGATGGGACTCAA 1

RESULT 16  
US-09-245-041-48  
; Sequence 48, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Moore, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 48  
; LENGTH: 24  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
US-09-245-041-48

Query Match 0.6%; Score 24; DB 3; Length 24;  
Best Local Similarity 100.0%; Pred. No. 0.099;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1631 GGACCAITCTTTAAGGACAGCCGAT 1654  
DB 1 GGACCAITCTTTAAGGACAGCCGAT 24

RESULT 17  
US-09-319-989-1/c  
; Sequence 1, Application US/09319989  
; Patent No. 6190914  
; GENERAL INFORMATION:  
; APPLICANT: Grivell, Leslie A.  
; APPLICANT: Teixeira De Mattos, Maarten J.  
; APPLICANT: Blom, Jolanda  
; TITLE OF INVENTION: METHODS FOR MODULATING METABOLIC PATHWAYS OF  
; TITLE OF INVENTION: MICRO-ORGANISMS AND MICRO-ORGANISMS OBTAINABLE BY SAID  
; TITLE OF INVENTION: METHODS  
; FILE REFERENCE: 24615-20123.00  
; CURRENT APPLICATION NUMBER: US/09/319,989  
; CURRENT FILING DATE: 1999-06-14  
; EARLIER APPLICATION NUMBER: PCT/NL97/00688  
; EARLIER FILING DATE: 1997-12-12  
; EARLIER APPLICATION NUMBER: EPO 96203520  
; EARLIER FILING DATE: 1996-12-12  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 1198  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: DNA construct  
; OTHER INFORMATION: encoding HAP2  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (201)..(995)

US-09-319-989-1

Query Match 0.6%; Score 24; DB 3; Length 1198;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGTTGCT 203  
DB 591 GCTGCTGCTGCTGCTGTTGCT 568

RESULT 18  
US-09-086-663A-70/c  
; Sequence 70, Application US/09086663A  
; Patent No. 6518063  
; GENERAL INFORMATION:  
; APPLICANT: DUCY, PATRICIA  
; APPLICANT: KARENTHY, GERARD  
; TITLE OF INVENTION: OSF2/CBFA1 COMPOSITIONS AND METHODS OF USE  
; FILE REFERENCE: UTSC:525  
; CURRENT APPLICATION NUMBER: US/09/086,663A  
; CURRENT FILING DATE: 1998-05-29  
; PRIOR APPLICATION NUMBER: 60/080,189  
; PRIOR FILING DATE: 1998-03-24  
; PRIOR APPLICATION NUMBER: 60/048,430  
; PRIOR FILING DATE: 1997-05-29  
; NUMBER OF SEQ ID NOS: 83  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 70  
; LENGTH: 2294  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1644)  
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
; OTHER INFORMATION: Primer  
US-09-086-663A-70

Query Match 0.6%; Score 24; DB 4; Length 2294;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGTTGCT 203  
DB 223 GCTGCTGCTGCTGCTGTTGCT 200

RESULT 19  
US-08-997-685A-1/c  
; Sequence 1, Application US/08997685A  
; Patent No. 6551821  
; GENERAL INFORMATION:  
; APPLICANT: The Trustees of Columbia University  
; APPLICANT: Kandel, Eric  
; TITLE OF INVENTION: Brain Cyclic Nucleotide Gated Ion Channel and Uses Thereof  
; FILE REFERENCE: 0575/54806  
; CURRENT APPLICATION NUMBER: US/08/997,685A  
; CURRENT FILING DATE: 1997-12-12  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 2733  
; TYPE: DNA  
; ORGANISM: mouse  
US-08-997-685A-1

Query Match 0.6%; Score 24; DB 4; Length 2733;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGTTGCT 203

```
Db      2281 GCTGCTGCTGCTGCTGCTGCTTGTGCT 2258
RESULT 20
US-09-086-663A-1/c
; Sequence 1, Application US/09086663A
; Patent No. 6518063
; GENERAL INFORMATION:
; APPLICANT: DUCY, PATRICIA
; APPLICANT: KARSENTY, GERARD
; TITLE OF INVENTION: OSP2/CBFA1 COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: UTSC:525
; CURRENT APPLICATION NUMBER: US/09/086,663A
; CURRENT FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/080,189
; PRIOR FILING DATE: 1998-03-24
; PRIOR APPLICATION NUMBER: 60/048,430
; PRIOR FILING DATE: 1997-05-29
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3334
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Primer
US-09-086-663A-1
Query Match      0.6%; Score 24; DB 4; Length 3334;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 180 GCTGCTGCTGCTGCTGCTGCTGCT 203
Db 531 GCTGCTGCTGCTGCTGCTGCTGCT 508

RESULT 21
US-09-245-041-24
; Sequence 24, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial Sequence
US-09-245-041-24
Query Match      0.5%; Score 23; DB 3; Length 29;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 GCCAAATAGATATAGAGACTTC 517
Db 7 GCCAAATAGATATAGAGACTTC 29

RESULT 22
US-09-245-041-7/c
; Sequence 7, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 221
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-7
Query Match      0.5%; Score 23; DB 3; Length 221;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2713 AAGGCTGCAACCTGCATCAACCC 2735
Db 118 AAGGCTGCAACCTGCATCAACCC 96

RESULT 23
US-09-245-041-6/c
; Sequence 6, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 226
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-6
Query Match      0.5%; Score 23; DB 3; Length 226;
Best Local Similarity 100.0%; Pred. No. 0.34;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2713 AAGGCTGCAACCTGCATCAACCC 2735
Db 123 AAGGCTGCAACCTGCATCAACCC 101

RESULT 24
US-08-684-101-1/c
; Sequence 1, Application US/08684101
; Patent No. 5846789
; GENERAL INFORMATION:
; APPLICANT: Abrams, John M.
; TITLE OF INVENTION: VERTEBRATE APOPTOSIS GENE:
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 2
```



FILE REFERENCE: GENSET.054PR2  
CURRENT APPLICATION NUMBER: US/09/621,976  
CURRENT FILING DATE: 2000-07-21  
NUMBER OF SEQ ID NOS: 19335  
SOFTWARE: Patent.pm  
SEQ ID NO 7843  
LENGTH: 365  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-621-976-7843

Query Match 0.5%; Score 22; DB 4; Length 365;

Best Local Similarity 100.0%; Pred. No. 1.1;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 246 GGCGGGGGGGGGGGGGGGGG 267

Db 119 GGCGGGGGGGGGGGGGGGGG 140

#### RESULT 28

US-09-148-545-36  
Sequence 36, Application US/09148545  
Patent No. 6590075  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: 70 Human Secreted Proteins  
FILE REFERENCE: PZ001P1  
CURRENT APPLICATION NUMBER: US/09/148,545  
CURRENT FILING DATE: 1998-09-04  
EARLIER APPLICATION NUMBER: PCT/US98/04482  
EARLIER FILING DATE: 1998-03-06  
EARLIER APPLICATION NUMBER: 60/040,162  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,333  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/038,621  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,161  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,626  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,334  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,336  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/040,163  
EARLIER FILING DATE: 1997-03-07  
EARLIER APPLICATION NUMBER: 60/047,615  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,600  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,597  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,502  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,633  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,583  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,617  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,618  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,503  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,592  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,581  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,584  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,500

EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,587  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,492  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,598  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,613  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,582  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,596  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,612  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,632  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,601  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,580  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,568  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,314  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,569  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,311  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,671  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,674  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,669  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,312  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,313  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,672  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,315  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/056,886  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,877  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,889  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,893  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,630  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,878  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,662  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,872  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,882  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,637  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,903  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,888  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,879  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,880  
EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,894  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,911  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,636  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,874  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,910  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,864  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,631  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,845  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,892  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/047,595  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/057,761  
EARLIER FILING DATE: 05-Sep-1997  
EARLIER APPLICATION NUMBER: 60/047,599  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,588  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,585  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,586  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,590  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,594  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,589  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,593  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/047,614  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,578  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/043,576  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/047,501  
EARLIER FILING DATE: 1997-05-23  
EARLIER APPLICATION NUMBER: 60/043,670  
EARLIER FILING DATE: 1997-04-11  
EARLIER APPLICATION NUMBER: 60/056,632  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,664  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,876  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,881  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,909  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,875  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,862  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,887  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/056,908  
EARLIER FILING DATE: 1997-08-22  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/057,650  
EARLIER FILING DATE: 1997-09-05  
EARLIER APPLICATION NUMBER: 60/056,884  
EARLIER FILING DATE: 1997-08-22  
NUMBER OF SEQ ID NOS: 280

SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 36  
LENGTH: 459  
Query Match 0.5%; Score 22; DB 4; Length 459;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 211 CCCTGCTGCTGCTGCTGCTG 232  
DB 34 CCCTGCTGCTGCTGCTGCTG 55  
RESULT 29  
US-08-998-416-915/C  
Sequence 915, Application US/08998416  
Patent No. 6239264  
GENERAL INFORMATION:  
APPLICANT: Philippsen, Peter  
APPLICANT: Pohlmann, Rainer  
APPLICANT: Steiner, Sabine  
APPLICANT: Mohr, Christine  
APPLICANT: Wendland, Jurgен  
APPLICANT: Knechtle, Philipp  
APPLICANT: Rebeschung, Corinne  
TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPHII  
TITLE OF INVENTION: AND USES THEREOF  
NUMBER OF SEQUENCES: 1152  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: No. 6239264artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: No. 6239264th Carolina  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/998,416  
FILING DATE: 24-DEC-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: CH 0016/97  
FILING DATE: 31-DEC-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Meigs, J. Timothy  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
TELEFAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 915:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 688 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
ORIGINAL SOURCE:  
ORGANISM: PAG1570RP  
US-08-998-416-915  
Query Match 0.5%; Score 22; DB 3; Length 688;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 180 GCTGCTGCTGCTGCTGCTG 201  
DB 652 GCTGCTGCTGCTGCTGCTG 631

RESULT 30  
US-08-195-152-1/c  
Sequence 1, Application US/08195152  
Patent No. 5679541  
GENERAL INFORMATION:  
APPLICANT: Bonini, Nancy M.  
APPLICANT: Leiserson, William M.  
APPLICANT: Benzer, Seymour  
TITLE OF INVENTION: PROGRAMMED CELL DEATH ANTAGONIST  
TITLE OF INVENTION: PROTEINS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/195,152  
FILING DATE: 14-FEB-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-59551/RPT/RMS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 1:  
FILING DATE: 14-FEB-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-59551/RPT/RMS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3231 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
US-08-195-152-1

Query Match 0.5%; Score 22; DB 1; Length 3231;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 TGCTGCTGCTGCTGTTGTGCT 203  
|||||  
Db 580 TGCTGCTGCTGCTGTTGTGCT 559

RESULT 31  
US-08-754-311B-1/c  
Sequence 1, Application US/08754311B  
Patent No. 6623937  
GENERAL INFORMATION:  
APPLICANT: Bonini, Nancy M.  
APPLICANT: Leiserson, William M.  
APPLICANT: Benzer, Seymour  
TITLE OF INVENTION: PROGRAMMED CELL DEATH ANTAGONIST  
TITLE OF INVENTION: PROTEINS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/754,311B  
FILING DATE: 21-NOV-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/195,152  
FILING DATE: 14-FEB-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Trecartin, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-59551/RPT/RMS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3231 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
US-08-754-311B-1

Query Match 0.5%; Score 22; DB 4; Length 3231;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 TGCTGCTGCTGCTGTTGTGCT 203  
|||||  
Db 580 TGCTGCTGCTGCTGTTGTGCT 559

RESULT 32  
US-08-785-310A-3  
Sequence 3, Application US/08785310A  
Patent No. 5840532  
GENERAL INFORMATION:  
APPLICANT: McKnight, Steven L.  
APPLICANT: Russell, David W.  
TITLE OF INVENTION: Neuronal PAS Domain Protein  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/785,310A  
FILING DATE: 21-JAN-1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: OSWAN, RICHARD A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UTSD:1226  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4010 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear



```
; MOLECULE TYPE: cDNA
US-08-785-310A-3

Query Match      0.5%; Score 22; DB 2; Length 4010;
Best Local Similarity 100.0%; Pred.No.1.2; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 247 GCCGCGCGCGCGCGCGCGCGCG 268
DB 94 GCCGCGCGCGCGCGCGCGCGCG 115

RESULT 33
US-09-491-356C-7/c
; Sequence 7, Application US/09491356C
; Patent No. 6566061
; GENERAL INFORMATION:
; APPLICANT: Philibert, Robert A.
; APPLICANT: Gims, Edward I.
; APPLICANT: Delisi, Lynn
; TITLE OF INVENTION: IDENTIFICATION OF POLYMORPHISMS IN THE PCTG4 REGION OF XQ13
; FILE REFERENCE: 9465.6US11
; CURRENT APPLICATION NUMBER: US/09/491.356C
; CURRENT FILING DATE: 2000-01-26
; PRIOR APPLICATION NUMBER: PCT/US99/09365
; PRIOR FILING DATE: 1999-04-29
; PRIOR APPLICATION NUMBER: 60/083,465
; PRIOR FILING DATE: 1998-04-29
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 6558
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-491-356C-7

Query Match      0.5%; Score 22; DB 4; Length 6558;
Best Local Similarity 100.0%; Pred.No.1.3; 0; Indels 0; Gaps 0;
Matches 22; Conservative 0; Mismatches 0;

QY 180 GCTGCTGCTGCTGCTGTTGTTG 201
DB 6184 GCTGCTGCTGCTGCTGTTGTTG 6163

RESULT 34
US-09-043-303-13/c
; Sequence 13, Application US/09043303
; Patent No. 6251589
; GENERAL INFORMATION:
; APPLICANT: TSUJI, Shoji
; APPLICANT: SANPEI, Kazujiro
; TITLE OF INVENTION: Method for Diagnosing Spinocerebellar Alaxia Type 2 and
; FILE REFERENCE: 0760-0241P
; CURRENT APPLICATION NUMBER: US/09/043.303
; CURRENT FILING DATE: 1998-05-18
; EARLIER APPLICATION NUMBER: PCT/JP96/01999
; EARLIER FILING DATE: 1996-07-18
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 69
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-043-303-13

Query Match      0.5%; Score 21; DB 3; Length 69;
Best Local Similarity 100.0%; Pred.No.3.2; 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 178 CGGCTGCTGCTGCTGCTGTTG 198
|||||

; MOLECULE TYPE: cDNA
US-09-621-976-8302/c
; Sequence 8302, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8302
; LENGTH: 168
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-8302

Query Match      0.5%; Score 21; DB 4; Length 168;
Best Local Similarity 100.0%; Pred.No.3.3; 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 248 CCGCGCGCGCGCGCGCGCGCG 268
DB 36 CCGCGCGCGCGCGCGCGCGCG 16

RESULT 36
US-09-621-976-12134
; Sequence 12134, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 12134
; LENGTH: 205
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-12134

Query Match      0.5%; Score 21; DB 4; Length 205;
Best Local Similarity 100.0%; Pred.No.3.4; 0; Indels 0; Gaps 0;
Matches 21; Conservative 0; Mismatches 0;

QY 250 GCGCGCGCGCGCGCGCGCGTG 270
DB 9 GCGCGCGCGCGCGCGCGCGTG 29

RESULT 37
US-09-621-976-17545/c
; Sequence 17545, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
```

```
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 17545
; LENGTH: 257
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-17545

Query Match          0.5%; Score 21; DB 4; Length 257;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 GCGCGCGCGCGCGGTGCGG 275
      |||||
Db 96 GCGCGCGCGCGGTGCGG 76

RESULT 38
US-09-370-838-145
; Sequence 145, Application US/09370838
; Patent No. 644425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 145
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapien.
US-09-370-838-145

Query Match          0.5%; Score 21; DB 4; Length 450;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGCGCGCGCGCGCGGTG 270
      |||||
Db 57 GCGCGCGCGCGCGCGGTG 77

RESULT 39
US-08-234-783-1/c
; Sequence 1, Application US/08234783
; Patent No. 5622835
; GENERAL INFORMATION:
; APPLICANT: Herlyn, Meenhard
; APPLICANT: Morris, Jennifer
; APPLICANT: Rauscher III, Frank J.
; APPLICANT: Rodeck, Ulrich
; TITLE OF INVENTION: Wt1 Monoclonal Antibodies and Methods of
; TITLE OF INVENTION: Use Therefor
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,907
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/234,783
; FILING DATE: 28-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST48AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 450 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..630
; US-08-234-783-1

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/234,783
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST48AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 633 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..630
; US-08-234-783-1

Query Match          0.5%; Score 21; DB 1; Length 633;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGCGCGCGCGCGCGGTG 270
      |||||
Db 229 GCGCGCGCGCGCGCGGTG 209

RESULT 40
US-08-456-907-1/c
; Sequence 1, Application US/08456907
; Patent No. 5633142
; GENERAL INFORMATION:
; APPLICANT: Herlyn, Meenhard
; APPLICANT: Morris, Jennifer
; APPLICANT: Rauscher III, Frank J.
; APPLICANT: Rodeck, Ulrich
; TITLE OF INVENTION: Wt1 Monoclonal Antibodies and Methods of
; TITLE OF INVENTION: Use Therefor
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howson and Howson
; STREET: Spring House Corporate Cntr, PO Box 457
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/456,907
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/234,783
; FILING DATE: 28-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: WST48AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
```

LENGTH: 633 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..630  
US-08-456-907-1

Query Match 0.5%; Score 21; DB 1; Length 633;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 250 GCGGCGGCGGCGGCGGCGGTG 270  
Db 229 GCGGCGGCGGCGGCGGCGGTG 209

## RESULT 41

PCT-US95-05523-1/c  
Sequence 1, Application PC/TUS9505523  
GENERAL INFORMATION:  
APPLICANT: The Wistar Institute of Anatomy and Biology  
TITLE OF INVENTION: W1 Monoclonal Antibodies and  
TITLE OF INVENTION: Methods of Use Therefor  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, PO Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: PCT/US95/05523

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/234,783  
FILING DATE: 28-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST48PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 633 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..630  
PCT-US95-05523-1

Query Match 0.5%; Score 21; DB 5; Length 633;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 250 GCGGCGGCGGCGGCGGCGGTG 270  
Db 229 GCGGCGGCGGCGGCGGCGGTG 209

## RESULT 42

US-08-835-099A-16  
Sequence 16, Application US/08835099A  
Patent No. 5874277  
GENERAL INFORMATION:  
APPLICANT: SHINTANI, Yasushi  
APPLICANT: NISHI, Kazuori  
APPLICANT: KAWAMOTO, Tomohiro  
TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION  
TITLE OF INVENTION: AND USE  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/835,099A  
FILING DATE: 04-APR-1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 083649/1996  
FILING DATE: 05-APR-1996  
APPLICATION NUMBER: 97105508.2  
FILING DATE: 03-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Resnick, David S  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 47342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX:  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 640 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-835-099A-16

Query Match 0.5%; Score 21; DB 2; Length 640;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 250 GCGGCGGCGGCGGCGGCGGTG 270  
Db 125 GCGGCGGCGGCGGCGGCGGTG 145

## RESULT 43

US-09-157-349-16  
Sequence 16, Application US/09157349  
Patent No. 6068990  
GENERAL INFORMATION:  
APPLICANT: SHINTANI, Yasushi  
APPLICANT: NISHI, Kazuori  
APPLICANT: KAWAMOTO, Tomohiro  
TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION  
TITLE OF INVENTION: AND USE  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP  
STREET: 130 Water Street

```
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/157,349
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,099
; FILING DATE:
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 804 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-835-099A-10
;
; Query Match 0.5%; Score 21; DB 3; Length 640;
; Best Local Similarity 100.0%; Pred. No. 3.6;
; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 250 GCGGCGGCGGCGGCGGCGGTG 270
; DB 125 GCGGCGGCGGCGGCGGCGGTG 145
;
; RESULT 44
; US-08-835-099A-10
; Sequence 10, Application US/08835099A
; Patent No. 5874277
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasushi
; APPLICANT: NISHI, Kazuori
; APPLICANT: KAWAMOTO, Tomohiro
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,099A
; FILING DATE: 04-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 083649/1996
;
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/157,349
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,099
; FILING DATE:
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 804 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-09-157-349-16
;
; Query Match 0.5%; Score 21; DB 2; Length 804;
; Best Local Similarity 100.0%; Pred. No. 3.6;
; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 250 GCGGCGGCGGCGGCGGCGGTG 270
; DB 125 GCGGCGGCGGCGGCGGCGGTG 145
;
; RESULT 45
; US-09-157-349-10
; Sequence 10, Application US/09157349
; Patent No. 6068990
; GENERAL INFORMATION:
; APPLICANT: SHINTANI, Yasushi
; APPLICANT: NISHI, Kazuori
; APPLICANT: KAWAMOTO, Tomohiro
; TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION
; TITLE OF INVENTION: AND USE
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/157,349
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,099
; FILING DATE:
; APPLICATION NUMBER: 97105508.2
; FILING DATE: 03-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Resnick, David S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 47342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX:
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 804 base pairs
; TYPE: nucleic acid
```

```
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-09-157-349-10

Query Match          0.5%  Score 21; DB 3; Length 804;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGCGGCGGCGGCGGCGGTG 270
    |||||
Db 125 GCGGCGGCGGCGGCGGCGGTG 145

RESULT 46
US-08-118-200-1
; Sequence 1, Application US/08118200
; Patent No. 6197500
; GENERAL INFORMATION:
; APPLICANT: SUTHERLAND, Grant R
; APPLICANT: RICHARDS, Robert I
; APPLICANT: SCHLESSINGER, David
; APPLICANT: NAGARAJA, Ramiah
; APPLICANT: KREMER, Eric J
; APPLICANT: YU, Sui
; APPLICANT: BAKER, Elizabeth
; APPLICANT: MULLEY, John C
; APPLICANT: MANDEL, Jean-Louis
; APPLICANT: PRITCHARD, Melanie April
; APPLICANT: LYNCH, Michael
; TITLE OF INVENTION: DNA SEQUENCES RELATED TO ISOLATED
; TITLE OF INVENTION: FRAGILE X SYNDROME
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,200
; FILING DATE: 09-SEP-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/802,650
; FILING DATE: 05-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,232
; FILING DATE: 20-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/638,518
; FILING DATE: 04-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/966,517
; FILING DATE: 23-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 020160-164
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1028 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-118-200-1

Query Match          0.5%  Score 21; DB 3; Length 1028;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGCGGCGGCGGCGGCGGTG 270
    |||||
Db 390 GCGGCGGCGGCGGCGGCGGTG 410

RESULT 47
US-08-458-745-1
; Sequence 1, Application US/08458745
; Patent No. 6242576
; GENERAL INFORMATION:
; APPLICANT: SUTHERLAND, Grant R
; APPLICANT: RICHARDS, Robert I
; APPLICANT: SCHLESSINGER, David
; APPLICANT: NAGARAJA, Ramiah
; APPLICANT: KREMER, Eric J
; APPLICANT: YU, Sui
; APPLICANT: BAKER, Elizabeth
; APPLICANT: MULLEY, John C
; APPLICANT: MANDEL, Jean-Louis
; APPLICANT: PRITCHARD, Melanie April
; APPLICANT: LYNCH, Michael
; TITLE OF INVENTION: DNA SEQUENCES RELATED TO ISOLATED
; TITLE OF INVENTION: FRAGILE X SYNDROME
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,745
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/118,200
; FILING DATE: 09-SEP-1993
; APPLICATION NUMBER: US 07/802,650
; FILING DATE: 05-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/672,232
; FILING DATE: 20-MAR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/638,518
; FILING DATE: 04-JAN-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/966,517
; FILING DATE: 23-DEC-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Crane-Feury, Sharon E
; REGISTRATION NUMBER: 36,113
; REFERENCE/DOCKET NUMBER: 020160-164
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1028 base pairs
; TYPE: nucleic acid
```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-458-745-1

Query Match          0.5%; Score 21; DB 3; Length 1028;
Best Local Similarity 100.0%; Pred. No. 3.6; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;

QY 250 GCGGCGGCGGCGGCGGCGGTG 270
Db 390 GCGGCGGCGGCGGCGGCGGTG 410

RESULT 48
US-08-765-875-1
; Sequence 1, Application US/08765875
; Patent No. 5914234
; GENERAL INFORMATION:
; APPLICANT: LEE, SE-JIN
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-11
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JURAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: US
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/765,875
; FILING DATE:
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/706,958
; FILING DATE:
; FILING DATE:
; FILING DATE: 08-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3641
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1393 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: HUMAN GDF-11
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 54..1274
US-08-765-875-1

Query Match          0.5%; Score 21; DB 2; Length 1393;
Best Local Similarity 100.0%; Pred. No. 3.7; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;

QY 248 CCGCGCGGCGGCGGCGGCGG 268
Db 136 CCGCGCGGCGGCGGCGGCGG 156

RESULT 49
US-08-795-671-1
; Sequence 1, Application US/08795671
; Patent No. 6008434
; GENERAL INFORMATION:
; APPLICANT: Se-Jin Lee and Alexandra McPherron
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-11
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/795,671
; FILING DATE: February 6, 1997
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 07265/106001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1393 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: HUMAN GDF-11
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 54..1274
US-08-795-671-1

Query Match          0.5%; Score 21; DB 3; Length 1393;
Best Local Similarity 100.0%; Pred. No. 3.7; Mismatches 0; Indels 0; Gaps 0;
Matches 21; Conservative 0;

QY 248 CCGCGCGGCGGCGGCGGCGG 268
Db 136 CCGCGCGGCGGCGGCGGCGG 156

RESULT 50
US-09-454-540-1
; Sequence 1, Application US/09454540
; Patent No. 6517835
; GENERAL INFORMATION:
; APPLICANT: Se-Jin Lee and Alexandra McPherron
; TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR-11
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; 
```

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/454,540  
FILING DATE: 06-DEC-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/795,671  
FILING DATE: February 6, 1997  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: HAILE, PH.D., LISA A.  
REGISTRATION NUMBER: 38,347  
REFERENCE/DOCKET NUMBER: 07265/106001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1393 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
IMMEDIATE SOURCE:  
CLONE: HUMAN GDF-11  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 54..1274  
US-09-454-540-1

Query Match 0.5%; Score 21; DB 4; Length 1393;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGCGCGCGCGG 268  
DB 136 CCGCGCGCGCGCGCGCGCGG 156

RESULT 51  
US-09-626-896-24  
Sequence 24, Application US/09626896  
Patent No. 6656475  
GENERAL INFORMATION:  
APPLICANT: Lee, Se-Jin  
APPLICANT: McPherron, Alexandra C.  
TITLE OF INVENTION: GROWTH DIFFERENTIATION FACTOR RECEPTORS.  
TITLE OF INVENTION: AGONISTS AND ANTAGONISTS THEREOF, AND METHODS OF USING SAME  
FILE REFERENCE: JHU1470-2  
CURRENT APPLICATION NUMBER: US/09/626,896  
CURRENT FILING DATE: 2000-07-27  
PRIOR APPLICATION NUMBER: 09/485,046  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: PCT/US98/15598  
PRIOR FILING DATE: 1998-07-28  
PRIOR APPLICATION NUMBER: 60/054,461  
PRIOR FILING DATE: 1997-08-01  
NUMBER OF SEQ ID NOS: 29  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 24  
LENGTH: 1393  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (54)...(1274)  
OTHER INFORMATION: GDF-11  
US-09-626-896-24

Query Match 0.5%; Score 21; DB 4; Length 1393;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGCGCGCGCGG 268  
DB 136 CCGCGCGCGCGCGCGCGCGG 156  
RESULT 52  
US-09-220-132-82  
Sequence 82, Application US/09220132  
Patent No. 6506607  
GENERAL INFORMATION:  
APPLICANT: Shyjan, Andrew W.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER  
FILE REFERENCE: 07334-074001  
CURRENT APPLICATION NUMBER: US/09/220,132  
CURRENT FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: US 60/079,303  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: US 60/068,821  
PRIOR FILING DATE: 1997-12-24  
NUMBER OF SEQ ID NOS: 191  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 82  
LENGTH: 1682  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-220-132-82

Query Match 0.5%; Score 21; DB 4; Length 1682;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGCGCGCGCGCGCGCGGTG 270  
DB 746 GCGCGCGCGCGCGCGCGGTG 766

RESULT 53  
US-08-857-636-58  
Sequence 58, Application US/08857636  
Patent No. 6552181  
GENERAL INFORMATION:  
APPLICANT: Dean, Michael Carlton  
APPLICANT: Hahn, Heidi Eve  
APPLICANT: Wicking, Carol  
APPLICANT: Christiansen, Jeffrey  
APPLICANT: Zaphitopoulos, Peter G.  
APPLICANT: Gallani, Mae R.  
APPLICANT: Shanley, Susan Mary  
APPLICANT: Chidambaram, Abirami  
APPLICANT: Vorechovsky, Igor  
APPLICANT: Holmberg-Lindstrom, Erika  
APPLICANT: Unden, Anne Birgitte  
APPLICANT: Gillies, Susan Alana  
APPLICANT: Negus, Kylie  
APPLICANT: Smyth, Ian Mcleod  
APPLICANT: Pressman, Carol Leah  
APPLICANT: Lefell, David J.  
APPLICANT: Gerrard, Bernard  
APPLICANT: Goldstein, Alisa Miriam  
APPLICANT: Mainwright, Brandon  
APPLICANT: Toftgard, Rune Carl-Magnus  
APPLICANT: Chenevix-Trench, Georgia  
APPLICANT: Bale, Allen E.  
TITLE OF INVENTION: A Basal Cell Carcinoma Tumor Suppressor Gene  
NUMBER OF SEQUENCES: 83  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
US-08-857-636-58





APPLICANT: Shyjan, Andrew W.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
FILE REFERENCE: 07334-074001  
CURRENT APPLICATION NUMBER: US/09/220,132  
CURRENT FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: US 60/079,303  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: US 60/068,821  
PRIOR FILING DATE: 1997-12-24  
NUMBER OF SEQ ID NOS: 191  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11  
LENGTH: 1860  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-220-132-11

Query Match 0.5%; Score 21; DB 4; Length 1860;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 GCGCGCGCGCGGTGTCTCGG 275  
DB 93 GCGCGCGCGCGGTGTCTCGG 73

## RESULT 56

US-09-220-132-67/c  
Sequence 67, Application US/09220132  
Patent No. 6506607  
GENERAL INFORMATION:  
APPLICANT: Shyjan, Andrew W.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
FILE REFERENCE: 07334-074001  
CURRENT APPLICATION NUMBER: US/09/220,132  
CURRENT FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: US 60/079,303  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: US 60/068,821  
PRIOR FILING DATE: 1997-12-24  
NUMBER OF SEQ ID NOS: 191  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 67  
LENGTH: 1860  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-220-132-67

Query Match 0.5%; Score 21; DB 4; Length 1860;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 GCGCGCGCGCGGTGTCTCGG 275  
DB 93 GCGCGCGCGCGGTGTCTCGG 73

## RESULT 57

US-09-220-132-137/c  
Sequence 137, Application US/09220132  
Patent No. 6506607  
GENERAL INFORMATION:  
APPLICANT: Shyjan, Andrew W.  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
FILE REFERENCE: 07334-074001  
CURRENT APPLICATION NUMBER: US/09/220,132  
CURRENT FILING DATE: 1998-12-23  
PRIOR APPLICATION NUMBER: US 60/079,303  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: US 60/068,821

PRIOR FILING DATE: 1997-12-24  
NUMBER OF SEQ ID NOS: 191  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 137  
LENGTH: 1860  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-220-132-137

Query Match 0.5%; Score 21; DB 4; Length 1860;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 GCGCGCGCGCGGTGTCTCGG 275  
DB 93 GCGCGCGCGCGGTGTCTCGG 73

## RESULT 58

US-08-937-067-5  
Sequence 5, Application US/08937067  
Patent No. 6433155  
GENERAL INFORMATION:  
APPLICANT: Umansky, Samuil  
APPLICANT: Melkonian, Hovsep  
TITLE OF INVENTION: A FAMILY OF GENES ENCODING  
TITLE OF INVENTION: APOPTOSIS-RELATED PEPTIDES; PEPTIDES ENCODED THEREBY AND  
TITLE OF INVENTION: METHODS OF USE THEREOF  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/937,067  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Lehnhardt, Susan K.  
REGISTRATION NUMBER: 33,943  
REFERENCE/DOCKET NUMBER: 23647-20018.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 813-5600  
TELEFAX: (650) 494-0792  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1984 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 216..1166  
US-08-937-067-5

Query Match 0.5%; Score 21; DB 4; Length 1984;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGGTGTCTCGG 268  
DB 94 CCGCGCGCGCGGTGTCTCGG 114

RESULT 59  
US-07-952-800-1  
; Sequence 1, Application US/07952800  
; Patent No. 5403925  
; GENERAL INFORMATION:  
; APPLICANT: OZATO, KEIKO  
; TITLE OF INVENTION: A NEW MEMBER OF THE NUCLEAR HORMONE  
; TITLE OF INVENTION: RECEPTOR SUPERFAMILY AND A CDNA CLONE THEREOF  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: TOWNSEND AND TOWNSEND  
; STREET: 1 MARKET PLAZA, STEUART TOWER, SUITE 2000  
; CITY: SAN FRANCISCO  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/952,800  
; FILING DATE: 19920928  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Kenneth A  
; REGISTRATION NUMBER: 31,677  
; REFERENCE/DOCKET NUMBER: 15280-21-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2130 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: DAUDI  
; CELL TYPE: T cell lymphoma  
; CELL LINE: DAUDI  
; IMMEDIATE SOURCE:  
; LIBRARY: CDNA DAUDI  
; CLONE: DAUDI6  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 180..1778  
; FEATURE:  
; NAME/KEY: polyA signal  
; LOCATION: 2109..2121  
US-07-952-800-1  
Query Match 0.5%; Score 21; DB 1; Length 2130;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 250 GCGGCGGCGGCGGCGGTG 270  
Db 318 GCGGCGGCGGCGGCGGTG 338  
RESULT 60  
US-09-220-132-80/c  
; Sequence 80, Application US/09220132  
; Patent No. 6506607  
; GENERAL INFORMATION:  
; APPLICANT: Shv'jan, Andrew W.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT  
; TITLE OF INVENTION: OF PROSTATE CANCER THERAPIES AND THE DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 07334-074001  
; CURRENT APPLICATION NUMBER: US/09/220,132

; CURRENT FILING DATE: 1998-12-23  
; PRIOR APPLICATION NUMBER: US 60/079,303  
; PRIOR FILING DATE: 1998-03-25  
; PRIOR APPLICATION NUMBER: US 60/068,821  
; PRIOR FILING DATE: 1997-12-24  
; NUMBER OF SEQ ID NOS: 191  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 80  
; LENGTH: 2634  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-220-132-80  
Query Match 0.5%; Score 21; DB 4; Length 2634;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 248 CCGCGCGGCGGCGGCGG 268  
Db 25 CCGCGCGGCGGCGGCGG 5  
RESULT 61  
US-09-620-312D-713  
; Sequence 713, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunging  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghaast  
; APPLICANT: Dmanac, Radoje T.  
; TITLE OF INVENTION: NO. 6569662el Nucleic Acids and  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt FL\_genes Version 1.0  
; SEQ ID NO 713  
; LENGTH: 3169  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (116)..(2146)  
US-09-620-312D-713  
Query Match 0.5%; Score 21; DB 4; Length 3169;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 248 CCGCGCGGCGGCGGCGG 268  
Db 544 CCGCGCGGCGGCGGCGG 564

```
RESULT 62
US-09-799-875-6/c
; Sequence 6, Application US/09799875
; Patent No. 6638721
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. 6638721el Human Protein Kinases and Uses
; TITLE OF INVENTION: Therefor
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/09/799,875
; CURRENT FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/659,287
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 3609
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-799-875-6
Query Match 0.5%; Score 21; DB 4; Length 3609;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 213 GCTGTGCTGCTGCTGCTGCC 233
Db 2674 GCTGTGCTGCTGCTGCTGCC 2654

RESULT 63
US-09-620-312D-348/c
; Sequence 348, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: PC_FL_genes Version 1.0
; SEQ ID NO 348
; LENGTH: 4062
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS

; LOCATION: (25)..(2976)
US-09-620-312D-348
Query Match 0.5%; Score 21; DB 4; Length 4062;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 178 CGGCTGCTGCTGCTGCTGTTG 198
Db 2892 CGGCTGCTGCTGCTGCTGTTG 2872

RESULT 64
US-08-540-406-18
; Sequence 18, Application US/08540406
; Patent No. 5837538
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
; STREET: Four Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: CA
; COUNTRY: US
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/540,406
; FILING DATE: 06-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Bertram I
; REGISTRATION NUMBER: 20015
; REFERENCE/DOCKET NUMBER: a60190-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-781-1989
; TELEFAX: 415-398-3249
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-540-406-18
Query Match 0.5%; Score 21; DB 2; Length 5288;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 248 CCGCGCGCGCGCGCGCGCGG 268
Db 353 CCGCGCGCGCGCGCGCGCGG 373

RESULT 65
US-08-656-055-18
; Sequence 18, Application US/08656055
; Patent No. 6027882
; GENERAL INFORMATION:
; APPLICANT: SCOTT, MATHEW P
; APPLICANT: GOODRICH, LISA V
; APPLICANT: JOHNSON, RONALD L
; TITLE OF INVENTION: Patched Genes and their Use
; NUMBER OF SEQUENCES: 19
```

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/656,055
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/540,406
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5288 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-656-055-18

Query Match 0.5%; Score 21; DB 3; Length 5288;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGGCGCGCGCGCGCGCGG 268
Db 353 CCGCGGCGCGCGCGCGCGCGG 373

RESULT 66
US-09-954-668-18
Sequence 18, Application US/08954668
Patent No. 6172200
GENERAL INFORMATION:
APPLICANT: SCOTT, MATHEW P
APPLICANT: GOODRICH, LISA V
APPLICANT: JOHNSON, RONALD L
TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley, Hoag & Eliot
STREET: One Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/954,668
FILING DATE: 20-Oct-1997
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P
REGISTRATION NUMBER: 36709
```

```

REFERENCE/DOCKET NUMBER: SUV-003.06
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-832-1000
TELEFAX: 617-832-7000
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5288 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-954-668-18

Query Match 0.5%; Score 21; DB 3; Length 5288;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGGCGCGCGCGCGCGCGG 268
Db 353 CCGCGGCGCGCGCGCGCGCGG 373

RESULT 67
US-08-918-658-18
Sequence 18, Application US/08918658
Patent No. 6429354
GENERAL INFORMATION:
APPLICANT: SCOTT, MATHEW P
APPLICANT: GOODRICH, LISA V
APPLICANT: JOHNSON, RONALD L
TITLE OF INVENTION: Patched Genes and their Use
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert
STREET: Four Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: US
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,658
FILING DATE: 22-Aug-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/656,055
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/540,406
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20015
REFERENCE/DOCKET NUMBER: a60190-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5288 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-08-918-658-18

Query Match 0.5%; Score 21; DB 4; Length 5288;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 248 CCGCGCGCGCGCGCGCGG 268  
Db 353 CCGCGCGCGCGCGCGCGG 373

## RESULT 68

US-09-724-631-18  
; Sequence 18, Application US/09724631  
; Patent No. 6551782  
; GENERAL INFORMATION:  
; APPLICANT: SCOTT, MATHEW P  
; GOODRICH, LISA V  
; JOHNSON, RONALD L  
; TITLE OF INVENTION: Patched Genes and their Use  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/724,631  
; FILING DATE: 28-Nov-85  
; CLASSIFICATION: <Unknown>  
; PRIORITY INFORMATION:  
; APPLICATION NUMBER: US/08/656,055  
; FILING DATE: 1996-05-31  
; APPLICATION NUMBER: 08/540,406  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rowland, Bertram I  
; REGISTRATION NUMBER: 20015  
; REFERENCE/DOCKET NUMBER: a60190-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1389  
; TELEFAX: 415-398-3249  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5288 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-724-631-18

Query Match 0.5%; Score 21; DB 4; Length 5288;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 248 CCGCGCGCGCGCGCGCGG 268  
Db 353 CCGCGCGCGCGCGCGCGG 373

## RESULT 69

US-08-954-701A-18  
; Sequence 18, Application US/08954701A  
; Patent No. 6610507  
; GENERAL INFORMATION:  
; APPLICANT: SCOTT, MATHEW P  
; GOODRICH, LISA V  
; JOHNSON, RONALD L  
; TITLE OF INVENTION: Patched Genes and their Use  
; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley, Hoag & Eliot  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII(text)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/954,701A  
; FILING DATE: 20-OCT-1997  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Vincent, Matthew P.  
; REGISTRATION NUMBER: 36709  
; REFERENCE/DOCKET NUMBER: SUV-003.08  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5288 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-954-701A-18

Query Match 0.5%; Score 21; DB 4; Length 5288;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 248 CCGCGCGCGCGCGCGCGG 268  
Db 353 CCGCGCGCGCGCGCGCGG 373

## RESULT 70

PCT-US95-13233-18  
; Sequence 18, Application PC/TUS9513233  
; GENERAL INFORMATION:  
; APPLICANT: THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY  
; TITLE OF INVENTION: Patched Genes and their Use  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: US  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/13233  
; FILING DATE: 06-OCT-1990  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Rowland, Bertram I  
; REGISTRATION NUMBER: 20015  
; REFERENCE/DOCKET NUMBER: a60190-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-781-1389  
; TELEFAX: 415-398-3249  
; INFORMATION FOR SEQ ID NO: 18:  
; SEQUENCE CHARACTERISTICS:

```
;
; LENGTH: 5288 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
PCT-US95-13233-18

Query Match
Best Local Similarity 100.0%; Pred. No. 3.9; Length 5288;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 248 CCGCGCGCGCGCGCGCGCGCG 268
Db 353 CCGCGCGCGCGCGCGCGCGCG 373

RESULT 71
US-09-799-875-4/c
; Sequence 4, Application US/09799875
; Patent No. 6638721
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Williamson, Mark
; TITLE OF INVENTION: No. 6638721el Human Protein Kinases and Uses
; FILE OF INVENTION: Therefor
; FILE REFERENCE: 35800/209996
; CURRENT APPLICATION NUMBER: US/09/799,875
; CURRENT FILING DATE: 2001-03-06
; PRIOR APPLICATION NUMBER: 60/182,059
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: 09/655,287
; PRIOR FILING DATE: 2000-09-12
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 5983
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(3723)
US-09-799-875-4

Query Match
Best Local Similarity 100.0%; Pred. No. 3.9; Length 5983;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 213 GCTGTTGCTGCTGCTGCTGCC 233
Db 2788 GCTGTTGCTGCTGCTGCTGCC 2768

RESULT 72
US-08-457-273B-41/c
; Sequence 41, Application US/08457273B
; Patent No. 5849995
; GENERAL INFORMATION:
; APPLICANT: Hayden, Michael
; APPLICANT: Lin, Biaoyang
; APPLICANT: Nasir, Jamal
; TITLE OF INVENTION: Mouse Model for Huntington's Disease and
; TITLE OF INVENTION: Related DNA Sequences
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Virginia Bennett
; STREET: PO Box 37428
; CITY: Raleigh
; STATE: No. 5849995th Carolina
; COUNTRY: US
; ZIP: 27627
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,273B
; FILING DATE:
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Bennett, Virginia C.
; REGISTRATION NUMBER: 37,092
; REFERENCE/DOCKET NUMBER: 3477-85A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-854-1400
; TELEFAX: 919-854-1401
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10348 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-08-457-273B-41

Query Match
Best Local Similarity 100.0%; Pred. No. 4; Length 10348;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGCGCGCGCGCGCGCGGTG 270
Db 460 GCGGCGCGCGCGCGCGCGGTG 440

RESULT 73
US-08-556-419-13/c
; Sequence 13, Application US/08556419C
; Patent No. 6093549
; GENERAL INFORMATION:
; APPLICANT: Ross, Christopher
; APPLICANT: Li, Xiao-Jiang
; APPLICANT: Li, Shi-Hua
; APPLICANT: Sharp, Alan
; APPLICANT: Lanahan, Anthony
; APPLICANT: Worley, Paul
; APPLICANT: Snyder, Solomon
; TITLE OF INVENTION: Huntingtin-associated protein
; FILE REFERENCE: 01107.52271
; CURRENT APPLICATION NUMBER: US/08/556,419C
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 13
; LENGTH: 10348
; TYPE: DNA
; ORGANISM: Homo sapiens
US-08-556-419-13

Query Match
Best Local Similarity 100.0%; Pred. No. 4; Length 10348;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGCGCGCGCGCGCGCGGTG 270
Db 460 GCGGCGCGCGCGCGCGCGGTG 440

RESULT 74
US-09-041-886-14/c
; Sequence 14, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bradesen, Dale E.
; APPLICANT: Rabilzadeh, Sharroz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
```

;; TITLE OF INVENTION: Polypeptides and Methods of Use  
;; NUMBER OF SEQUENCES: 72  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Campbell & Flores LLP  
;; STREET: 4370 La Jolla Village Drive, Suite 700  
;; CITY: San Diego  
;; STATE: California  
;; COUNTRY: United States  
;; ZIP: 92122  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/041.886  
;; FILING DATE:  
;; CLASSIFICATION:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Campbell, Cathryn A.  
;; REGISTRATION NUMBER: 31.815  
;; REFERENCE/DOCKET NUMBER: P-LJ 2626  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 535-9001  
;; TELEFAX: (619) 535-8949  
;; INFORMATION FOR SEQ ID NO: 14:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10348 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: DNA (genomic)  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 316..9748  
US-09-041-886-14

Query Match 0.5%; Score 21; DB 3; Length 10348;  
Best Local Similarity 100.0%; Pred.No. 4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGGCGGGCGGGCGGGCGGTG 270  
|||||  
DB 460 GCGGGCGGGCGGGCGGGCGGTG 440

RESULT 75  
US-08-246-982A-5/c  
;; Sequence 5, Application US/08246982A  
;; Patent No. 5686288  
;; GENERAL INFORMATION:  
;; APPLICANT: MacDonald, Marcy E.  
;; APPLICANT: Ambrose, Christine M.  
;; APPLICANT: Duyao, Mabel P.  
;; APPLICANT: Gusella, James F.  
;; TITLE OF INVENTION: Huntingtin DNA, Protein And Uses Thereof  
;; NUMBER OF SEQUENCES: 25  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
;; STREET: 1100 New York Avenue  
;; CITY: Washington  
;; STATE: D.C.  
;; COUNTRY: U.S.A.  
;; ZIP: 20005  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/246,982A  
;; FILING DATE: May 20, 1994  
;; CLASSIFICATION: 435

;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Goldstein, Jorge, A.  
;; REGISTRATION NUMBER: 29,021  
;; REFERENCE/DOCKET NUMBER: 0609.3880002  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (202) 371-2600  
;; TELEFAX: (202) 371-2540  
;; INFORMATION FOR SEQ ID NO: 5:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 10366 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: 316..9748  
US-08-246-982A-5

Query Match 0.5%; Score 21; DB 1; Length 10366;  
Best Local Similarity 100.0%; Pred.No. 4;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 250 GCGGGCGGGCGGGCGGGCGGTG 270  
|||||  
DB 460 GCGGGCGGGCGGGCGGGCGGTG 440

Search completed: March 1, 2004, 20:50:42  
Job time : 328 secs

***This Page Blank (uspto)***





[illegible]

QY 1441 GATTGGATAGAACACATGGAGTATATTACACACCAGGGTGCCTTGTGCAAGGGGT 1500  
DB 1441 GATTGGATAGAACACATGGAGTATATTACACACCAGGGTGCCTTGTGCAAGGGGT 1500  
QY 1501 TACGGCCATAGCAGTGTTCAGACCATAGACACAGGCGCCCTATACGTTTCATGTTGGCTAC 1560  
DB 1501 TACGGCCATAGCAGTGTTCAGACCATAGACACAGGCGCCCTATACGTTTCATGTTGGCTAC 1560  
QY 1561 AAGGCTTTTCAGTGCCAAATAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGGAT 1620  
DB 1561 AAGGCTTTTCAGTGCCAAATAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGGAT 1620  
QY 1621 ACCAGATGTGGACATCTTAAAGACAGCCGATTTTCGTTACTTGTGACACAGCTGTG 1680  
DB 1621 ACCAGATGTGGACATCTTAAAGACAGCCGATTTTCGTTACTTGTGACACAGCTGTG 1680  
QY 1681 ATAGTGTAGTGAACCATGCTGGTGTGGGGGAAACACACACATGACATCTATGAGC 1740  
DB 1681 ATAGTGTAGTGAACCATGCTGGTGTGGGGGAAACACACACATGACATCTATGAGC 1740  
QY 1741 CATGGCGCAAAATGCTCTCTTCAGATTTATGCGCTTATGACATGCTGTGACCGCTGG 1800  
DB 1741 CATGGCGCAAAATGCTCTCTTCAGATTTATGCGCTTATGACATGCTGTGACCGCTGG 1800  
QY 1801 TCAGTGTCTCCAGACCTGATCTCCACATGATGTCACAGATTTGGCCATTCAGCAGTC 1860  
DB 1801 TCAGTGTCTCCAGACCTGATCTCCACATGATGTCACAGATTTGGCCATTCAGCAGTC 1860  
QY 1861 TTACACACAGCACCATGATGTGTGTTGGTGGTTCATATAGTCTCTCCCTCAGCGACATC 1920  
DB 1861 TTACACACAGCACCATGATGTGTGTTGGTGGTTCATATAGTCTCTCCCTCAGCGACATC 1920  
QY 1921 CTGTTATTCAGCTGGAAACAGTGTGATGCGCATGCGAGTGAAGCGCTGTTTATGACGCA 1980  
DB 1921 CTGTTATTCAGCTGGAAACAGTGTGATGCGCATGCGAGTGAAGCGCTGTTTATGACGCA 1980  
QY 1981 GGACTGTGATTCGTTGTGTGAAACACAGGTCGTCTCAGTGTATCTCGTGGGCGCTG 2040  
DB 1981 GGACTGTGATTCGTTGTGTGAAACACAGGTCGTCTCAGTGTATCTCGTGGGCGCTG 2040  
QY 2041 GCAACTGTATGAACAGAGAAAGTAAATCAGATGTTTTTCCAAAGAACTCTTTGAC 2100  
DB 2041 GCAACTGTATGAACAGAGAAAGTAAATCAGATGTTTTTCCAAAGAACTCTTTGAC 2100  
QY 2101 CATGACAGATGACACACAGATGTTTACAGCTGTACAGCCACACCAATGACTGC 2160  
DB 2101 CATGACAGATGACACACAGATGTTTACAGCTGTACAGCCACACCAATGACTGC 2160  
QY 2161 CACTGGTCAATGACCATTTGTGTCACAGAACACACAGCTGCTCAGAGGCGCAGATCTCC 2220  
DB 2161 CACTGGTCAATGACCATTTGTGTCACAGAACACACAGCTGCTCAGAGGCGCAGATCTCC 2220  
QY 2221 ATTTTGTAGTATGAGAAATGCGCCAGGATTAACCCCATGTACTGTATACAGAGACC 2280  
DB 2221 ATTTTGTAGTATGAGAAATGCGCCAGGATTAACCCCATGTACTGTATACAGAGACC 2280  
QY 2281 AGCTCAGAGCTGTGCGCTGGACAGAACTGCGAGTGGAGCCCGGAATCAGAGTGC 2340  
DB 2281 AGCTCAGAGCTGTGCGCTGGACAGAACTGCGAGTGGAGCCCGGAATCAGAGTGC 2340  
QY 2341 ATTGCGCTGCCGAAATATCTGTGGCATTTGGTGGCAATTTGTTGAAATCTATGTTG 2400  
DB 2341 ATTGCGCTGCCGAAATATCTGTGGCATTTGGTGGCAATTTGTTGAAATCTATGTTG 2400  
QY 2401 AAAATTAATCTGCGCAAGGAAATATGACATGCTAAATGTTCTGTAGGAAACCAAT 2460  
DB 2401 AAAATTAATCTGCGCAAGGAAATATGACATGCTAAATGTTCTGTAGGAAACCAAT 2460  
QY 2461 GCCCTTTGGCTTCTCTTACACCCAGAGAGGTAGAAATTTGCTTTAAGCAGCTGCGA 2520  
DB 2461 GCCCTTTGGCTTCTCTTACACCCAGAGAGGTAGAAATTTGCTTTAAGCAGCTGCGA 2520  
QY 2521 ATAATGAGTATCTCAGAGCATGTCAGCTCACCTTAAACCCCATGGCTGGCTTCGG 2580

DB 2521 ATAATGAGTATCTCAGAGCATGTCAGCTCACCTTAAACCCCATGGCTGGCTTCGG 2580  
QY 2581 AAGATCAATGTCTCTTACTGGTGTGGGAGATATGTCCCATTTACAAATAGTTTACTA 2640  
DB 2581 AAGATCAATGTCTCTTACTGGTGTGGGAGATATGTCCCATTTACAAATAGTTTACTA 2640  
QY 2641 CAGTGGATCCGCTGTGAGCCAGTGTGATTTCTGTGGAATTTTATCAGAACCCAGT 2700  
DB 2641 CAGTGGATCCGCTGTGAGCCAGTGTGATTTCTGTGGAATTTTATCAGAACCCAGT 2700  
QY 2701 ACTCGGGGACTGAAGCTGCAACCTGCAACCCATCAACCTCAATGCTGTGTGAAAGG 2760  
DB 2701 ACTCGGGGACTGAAGCTGCAACCTGCAACCCATCAACCTCAATGCTGTGTGAAAGG 2760  
QY 2761 CTTGCAAAACACAGTGTCTAAGCAGTGTCCGAGCACCATGTGCTTGAAGCACAGTGTGA 2820  
DB 2761 CTTGCAAAACACAGTGTCTAAGCAGTGTCCGAGCACCATGTGCTTGAAGCACAGTGTGA 2820  
QY 2821 GATTGACACAGCGGAGCTCTGAGTGCATGTTGGTGCAGCAACATGAAGCAGTGTGTGAC 2880  
DB 2821 GATTGACACAGCGGAGCTCTGAGTGCATGTTGGTGCAGCAACATGAAGCAGTGTGTGAC 2880  
QY 2881 TCCAAATGCTATGTGGCCCTCTTCCCTTTTGGCCAGTGTATGGAATGATACGATGAC 2940  
DB 2881 TCCAAATGCTATGTGGCCCTCTTCCCTTTTGGCCAGTGTATGGAATGATACGATGAC 2940  
QY 2941 ACCTGCCCCCTGAAAATTTGTTTCCAGCTACTGTACCTGTAGTCAATTTGTTGAGCAACA 3000  
DB 2941 ACCTGCCCCCTGAAAATTTGTTTCCAGCTACTGTACCTGTAGTCAATTTGTTGAGCAACA 3000  
QY 3001 GGTGTGGTGTGTACTGATCCAGCAATCTGCAAGGAAATGCAATAGAGGTTCC 3060  
DB 3001 GGTGTGGTGTGTACTGATCCAGCAATCTGCAAGGAAATGCAATAGAGGTTCC 3060  
QY 3061 TATAAAGGACCAAGTGAAGTGTCTTCCAAAGCCCTTACAGGAAATTTCTATCCACAGCC 3120  
DB 3061 TATAAAGGACCAAGTGAAGTGTCTTCCAAAGCCCTTACAGGAAATTTCTATCCACAGCC 3120  
QY 3121 CTGCTCAATTCAGCATGTCTTAGAGACAGCAGATCAATGCTGCTTCTTACTCT 3180  
DB 3121 CTGCTCAATTCAGCATGTCTTAGAGACAGCAGATCAATGCTGCTTCTTACTCT 3180  
QY 3181 CCAGCTTGCAATGCAACGCGCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGT 3240  
DB 3181 CCAGCTTGCAATGCAACGCGCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGT 3240  
QY 3241 GAGAACCTGACACAGGCAAGCACTGCGAGACCTGCAATCTGCGTTCACGGTATCC 3300  
DB 3241 GAGAACCTGACACAGGCAAGCACTGCGAGACCTGCAATCTGCGTTCACGGTATCC 3300  
QY 3301 ACCAATGGAGGAAATGTGAGCCATGCAATGGGCAAGCTCTCTGTGCAACAC 3360  
DB 3301 ACCAATGGAGGAAATGTGAGCCATGCAATGGGCAAGCTCTCTGTGCAACAC 3360  
QY 3361 AACACGGGCAAGTGTCTTGTGACCAACAGGGCGTCAAGGGGACAGTGTGATGT 3420  
DB 3361 AACACGGGCAAGTGTCTTGTGACCAACAGGGCGTCAAGGGGACAGTGTGATGT 3420  
QY 3421 GAGGTAGAAATCGATACCAAGAAACCTCTCAGAGGAAATGTTATTTATCTCTTCT 3480  
DB 3421 GAGGTAGAAATCGATACCAAGAAACCTCTCAGAGGAAATGTTATTTATCTCTTCT 3480  
QY 3481 ATTGAGCTATCAGTTCACCTTTAGTCTATCCAGGAAAGATGATCGCTATTACAGCTATC 3540  
DB 3481 ATTGAGCTATCAGTTCACCTTTAGTCTATCCAGGAAAGATGATCGCTATTACAGCTATC 3540  
QY 3541 AATTTTGTGGTACTCTGACGAAACAAACAGGGATTTGACATGTTTCAATATGCTCC 3600  
DB 3541 AATTTTGTGGTACTCTGACGAAACAAACAGGGATTTGACATGTTTCAATATGCTCC 3600  
QY 3601 AAGATTTTCAACCTCAACATCACTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660





QY 3783 CTGGCCCATCAAAATTCAGGTGCAAACTGAAACAATGA 3819  
 Db 3780 CTGGCCCATCAAAATTCAGGTGCAAACTGAAACAATGA 3816

## RESULT 3

AR164816

LOCUS AR164816 8589 bp DNA linear PAT 17-OCT-2001

DEFINITION Sequence 14 from patent US 6274339.

ACCESSION AR164816

VERSION AR164816.1 GI:16238025

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 8589)

AUTHORS Moore, K. and Nagle, D. Lynn.

TITLE Methods and compositions for the diagnosis and treatment of body

weight disorders including obesity

JOURNAL Patent: US 6274339-A 14 14-AUG-2001;

FEATURES Location/Qualifiers

source 1..8589

/organism="unknown"

/mol\_type="unassigned DNA"

ORIGIN

Query Match 96.8%; Score 3696.2; DB 6; Length 8589;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 3744; Conservative 0; Mismatches 48; Indels 9; Gaps 1;

3 GGTGGCCGAGCGCGGCACTGAGCAAGGCTGAGGAGGAGCGCGGCGAGCGCAgC 62

9 GGTGGCCGAGCGCGGCACTGAGCAAGGCTGAGGAGGAGCGCGGCGAGCGCAgC 68

63 GTCGCGGCGAGGAGCGCGGCGCACTGAGCAAGGCTGAGGAGGAGCGCGGCGAGCGCAgC 122

69 GTCGCGGCGAGGAGCGCGGCGCGCACCGACCTGACCGCGAGCGGCGCTGGAGGCC 128

123 GGGGCTGGGGCGCGGCTGGCGCTCCCGGCGTGTCTCCACCGCTGGCGCGCGCGCGCT 182

129 GGGACCGCGCGCGCGCTGTGTCTCCCGGCGTGTGTCTCCCGGCGCGCGCGCGCGCG 188

183 GCTGCTGT 242

189 GCTGCTGT 239

243 CGAGCG 302

240 CGAGCG 299

303 TGACCG 362

300 TGACCG 359

363 CG 422

360 CG 419

423 TTCTGGGTTTGTGACAGATGACCTGGAAATTTATTAATACAAACGAAAGTGCAGTGGCT 482

420 TTCTGGGTTTGTGACAGATGACCTGGAAATTTATTAATACAAACGAAAGTGCAGTGGCT 479

483 CATTGAAGGACAGCGCAATAGAAATAGAGCTTGGTTTCAATCATTTTGTCTACAGAGTG 542

480 CATTGAAGGACAGCGCAATAGAAATAGAGCTTGGTTTCAATCATTTTGTCTACAGAGTG 539

543 TAGTTGGGACCAATTATATGTTTATGATGGGAGCTCAATTTATGCACCGCTAGTTGCTGC 602

540 TAGTTGGGACCAATTATATGTTTATGATGGGAGCTCAATTTATGCACCGCTAGTTGCTGC 599

603 ATTTAGTGGCCTCAITGTTTCTTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTC 662

600 ATTTAGTGGCCTCAITGTTTCTTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTC 659

QY 663 CACATCAGGTTATGCCTTGTCTGCAATTTTTTTAGTGTGCTGTATTAATTTGACTGGATT 722  
 Db 660 CACATCAGGTTATGCCTTGTCTGCAATTTTTTTAGTGTGCTGTATTAATTTGACTGGATT 719  
 QY 723 TAATATTACTTACAGTTTGTGATGTGCTCAATAAATCTGCTCAGGCCGAGGAGTGATA 782  
 Db 720 TAATATTACTTACAGTTTGTGATGTGCTCAATAAATCTGCTCAGGCCGAGGAGTGATA 779  
 QY 783 GATCAGTAATAGCAGCGAAACTCTTGAATGTGAATCTTCTGAAAACCTGAAAAGGTGAAC 842  
 Db 780 GATCAGTAATAGCAGCGATCTCTTGAATGTGAATCTTCTGAAAACCTGAAAAGGTGAAC 839  
 QY 843 ATGTGACATTCCTCCTGATGACAGCAAACTGTGTTTTTCTCATCGAGGATCTGCAATTC 902  
 Db 840 ATGTGACATTCCTCCTGATGACAGCAAACTGTGTTTTTCTCATCGAGGATCTGCAATTC 899  
 QY 903 AAGTGTATGTCAGAGGATGCTCCTGCTTCTCAGACTCGCAGGTCCTGAGTGTTCAGTTCC 962  
 Db 900 AAGTGTATGTCAGAGGATGCTCCTGCTTCTCAGACTCGCAGGTCCTGAGTGTTCAGTTCC 959  
 QY 963 TGTACCAGCTAACCCAGTCAATTTTGGACTCGAGAGGAATATTTAACTTAAAGCTCCCCAG 1022  
 Db 960 TGTACCAGCTAACCCAGTCAATTTTGGACTCGAGAGGAATATTTAACTTAAAGCTCCCCAG 1019  
 QY 1023 AGCATCTCAATAAGCTGTGCTCAATCGAAACATTTATGCGGTGTGTGGAGGATATATGTT 1082  
 Db 1020 AGCATCTCAATAAGCTGTGCTCAATCGAAACATTTATGCGGTGTGTGGAGGATATATGTT 1079  
 QY 1083 CAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGCTTCC 1142  
 Db 1080 CAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGCTTCC 1139  
 QY 1143 ACTAAACCGTTCTGTGAACAATGTGGTTGTGTAGATATGTCATTTCTTGGCATATACAA 1202  
 Db 1140 ACTAAACCGTTCTGTGAACAATGTGGTTGTGTAGATATGTCATTTCTTGGCATATACAA 1199  
 QY 1203 GGATAAAATTTACATGTATGGAGGAAATTTGATCAACTGGGAATGTACCAATGAGTT 1262  
 Db 1200 GGATAAAATTTACATGTATGGAGGAAATTTGATCAACTGGGAATGTACCAATGAGTT 1259  
 QY 1263 GAGAGTTTTTCAATTCATAATGAGTCATGGGTGTGTGTGACCCCTTAAGGCAAGAGCA 1322  
 Db 1260 GAGAGTTTTTCAATTCATAATGAGTCATGGGTGTGTGTGACCCCTTAAGGCAAGAGCA 1319  
 QY 1323 GTATGAGTGGTTGGGCACTCTGCAACATTTGTAACACTGGAAGATGGCCGAGTGGTCAT 1382  
 Db 1320 GTATGAGTGGTTGGGCACTCTGCAACATTTGTAACACTGGAAGATGGCCGAGTGGTCAT 1379  
 QY 1383 GCTGGTCACTTTTGGTCACTGCCCTCTCTATGGATATAAAGCAATGTGAGGAATATGA 1442  
 Db 1380 GCTGGTCACTTTTGGTCACTGCCCTCTCTATGGATATAAAGCAATGTGAGGAATATGA 1439  
 QY 1443 TTTGGATAGACACATGAGTATATTACACCCAGGTCGCCCTGTGCAAGGGGTTA 1502  
 Db 1440 TTTGGATAGACACATGAGTATATTACACCCAGGTCGCCCTGTGCAAGGGGTTA 1499  
 QY 1503 CGGCCATAGCAGTGTTTACGACCATAGGACAGGGGCCCTTATACGTTTCATGGTGGGTACAA 1562  
 Db 1500 CGGCCATAGCAGTGTTTACGACCATAGGACAGGGGCCCTTATACGTTTCATGGTGGGTACAA 1559  
 QY 1563 GGCCTTTCAGTGCATAAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATAC 1622  
 Db 1560 GGCCTTTCAGTGCATAAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATAC 1619  
 QY 1623 CCAGATGTGGACCATTTCTTAAGACAGCGGATTTTCCGTTACTTGGCACACAGCTGTGAT 1682  
 Db 1620 CCAGATGTGGACCATTTCTTAAGACAGCGGATTTTCCGTTACTTGGCACACAGCTGTGAT 1679  
 QY 1683 AGTGAAGTGAACCATGCTGGTGTGTGGGGGAAACACACAATGACATCTATGAGCCA 1742  
 Db 1680 AGTGAAGTGAACCATGCTGGTGTGTGGGGGAAACACACAATGACATCTATGAGCCA 1739

Qy	1743	TGGGCGCAAAAGTCTTCTTTTCAGATTTTCATGGCGCTATGACATTTGCCCTGTGACCGCTGGTCT	1802
Db	1740	TGGGCGCAAAATGTTTCTTTTCAGATTTTCATGGCGCTATGACATTTGCCCTGTGACCGCTGGTCT	1799
Qy	1803	AGTGCCTCCAGACACCTGATCTCTCCACCATGATGCAACAGATTTGGCCATTTACAGAGCTTT	1862
Db	1800	AGTGCCTCCAGACACCTGATCTCCACCATGATGCAACAGATTTGGCCATTTACAGAGCTTT	1859
Qy	1863	ACACAACAGCACCATGTATGTGTTCGGTGGTTTTCAATAGTCTCCTCCTCAGCGACATCCT	1922
Db	1860	ACACAACAGCACCATGTATGTGTTCGGTGGTTTTCAATAGTCTCCTCCTCAGCGACATCCT	1919
Qy	1923	GGTATTACCTCGGACACAGTGTATGGCATCGGAGTGAAGCCGTGTTTATAGCACGAG	1982
Db	1920	GGTATTACCTCGGACACAGTGTATGGCATCGGAGTGAAGCCGTGTTTATAGCACGAG	1979
Qy	1983	ACCTGGTATTCGGTGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGGCGCTGGC	2042
Db	1980	ACCTGGTATTCGGTGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGGCGCTGGC	2039
Qy	2043	AACGTATGAACACAGAGAAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCA	2102
Db	2040	AACGTATGAACACAGAGAAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCA	2099
Qy	2103	TGACAGATGTGACCCAGCACACAGATTTGTACAGCTGTACAGCCAAACCAATGACTGGCCA	2162
Db	2100	TGACAGATGTGACCCAGCACACAGATTTGTACAGCTGTACAGCCAAACCAATGACTGGCCA	2159
Qy	2163	CTGGTCAATGACCATTTGTGTCCCGAGGAACACACAGCTGTCAAGAGGCCAGATCTCCAT	2222
Db	2160	CTGGTCAATGACCATTTGTGTCCCGAGGAACACACAGCTGTCAAGAGGCCAGATCTCCAT	2219
Qy	2223	TTTTAGTATGAGAAATTCGCCCAAGGATTAACCCATGTACTACTGTAAACAAGAGACCAG	2282
Db	2220	TTTTAGTATGAGAAATTCGCCCAAGGATTAACCCATGTACTACTGTAAACAAGAGACCAG	2279
Qy	2283	CTGACGAGCTGTGCCCTGACACAGAACTGCCAGTGGGAGCCCGGAAATCAGGAGTGCAT	2342
Db	2280	CTGACGAGCTGTGCCCTGACACAGAACTGCCAGTGGGAGCCCGGAAATCAGGAGTGCAT	2339
Qy	2343	TGCCCTGCCGAAAATCTGTGGCATTTGCTGGCATTTGGTTGGAAATCATGTTTGA	2402
Db	2340	TGCCCTGCCGAAAATCTGTGGCATTTGCTGGCATTTGGTTGGAAATCATGTTTGA	2399
Qy	2403	AAATTACTTACGCCAAGGAGAAATTATGACAAATGCTAAATTTCTGTAGGAACACCAATGC	2462
Db	2400	AAATTACTTACGCCAAGGAGAAATTATGACAAATGCTAAATTTCTGTAGGAACACCAATGC	2459
Qy	2463	CTTTTGGCTTCTTTACAAACCCAGAGAGGTAGAAATTTGTCTTTAAGCAGTGGAAAT	2522
Db	2460	CTTTTGGCTTCTTTACAAACCCAGAGAGGTAGAAATTTGTCTTTAAGCAGTGGAAAT	2519
Qy	2523	AAATGCAGTATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTGGCCCTTCGGAA	2582
Db	2520	AAATGCAGTATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTGGCCCTTCGGAA	2579
Qy	2583	GATCAAAATGTGTCTTACTGTGGTGTGGGAAGATATGTCCCAATTTACAAATAGTTTACTACA	2642
Db	2580	GATCAAAATGTGTCTTACTGTGGTGTGGGAAGATATGTCCCAATTTACAAATAGTTTACTACA	2639
Qy	2643	GTGGATCCGCTGTAGCCGACAGTGTGATCTGGAAATTTTATCAGAACCCAGTAC	2702
Db	2640	GTGGATCCGCTGTAGCCGACAGTGTGATCTGGAAATTTTATCAGAACCCAGTAC	2699
Qy	2703	TCGGGGACTGAAGGCTCAACCTGCATCAACCCCACTCAATGGTAGTGTCTGTGAAAGGCC	2762
Db	2700	TCGGGGACTGAAGGCTCAACCTGCATCAACCCCACTCAATGGTAGTGTCTGTGAAAGGCC	2759
Qy	2763	TGGAAACACAGTCTTAAGCAGTCCCGGACACCAATGTGCTTGGAGGACAGATGTGGAGA	2822
Db	2760	TGGAAACACAGTCTTAAGCAGTCCCGGACACCAATGTGCTTGGAGGACAGATGTGGAGA	2819
Qy	2823	TTGACACAGGGGACGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGGACTC	2882

RESULT 4  
AF034957  
LOCUS  
DEFINITION



ACCESSION AF034957  
 VERSION AF034957.1 GI:3676346  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 3597)  
 Duke-Cohan,J.S., Gu,J., McLaughlin,D.F., Xu,Y., Freeman,G.J. and Schlozman,S.F.  
 TITLE Attractin (DPPT-L), a member of the CUB family of cell adhesion and guidance proteins, is secreted by activated human T lymphocytes and modulates immune cell interactions  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)  
 MEDLINE 98409658  
 PUBMED 9736737  
 REFERENCE 2 (bases 1 to 3597)  
 AUTHORS Duke-Cohan,J.S., Gu,J., Ao,Z., McLaughlin,D.F., Freeman,G.J. and Schlozman,S.F.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS, Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA  
 FEATURES  
 source  
 Location/Qualifiers  
 1..3597  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 1..3597  
 /gene="attractin"  
 1..3597  
 /gene="attractin"  
 /note="secreted T cell activation protein; Method: conceptual translation with partial peptide sequencing; DPPT-L"  
 /codon\_start=1  
 /product="Attractin"  
 /protein\_id="AAC61902.1"  
 /db\_xref="GI:3676347"  
 /translation="MVAAATEARLREARTAAALAGSGGPHVNGRCNPGTGC VCPAGVGEQCHCGFRLTGSSGFVTDGKNYKTKCTWLLIEGPNLMRLRPH FATECSWHLVYDGSIIYAPLVAAPSLIVPERDNETPEVVAATSGYALLHFFSDA AYNLGNITVYDFMCPNNSGRGECIKSNSDVECESENWKEACDIPHCYDNG FPHRGICNSDVRCSCFSDWQPGSCVFPANQSFRTREYSNLKLPRAHKAIVNG NIMVVGGMNHSYNNMLAYDLASREWLPLNRSVNNVVVRYGHSILYKDIYMYG KIDISTGNVWELRVFTHNESWLLTPKAQVAVVCHSAHIVTLKNGRVMLVIFG HCPVGLYSNVQEVLDKNTWLLHTQCALVQGGYGHSSVVDHRTALYVHGKYAS ANKYLADLLRYVDQWMTLLKDSREFRYLHTAVISGTMVLFVGGNTHNDTMSHG ACPSDFMAYDIADRSVLPRPDSITWMSDILPAVLHNSMTWVFGFNLSLSDI LFTSEQDAHRSEAAALAGGIRCVNWTSSQCI SWALADEQEBKLKSESKRT LDHRCQDHTCYSTANTNDCHNDRHVPNHSCEGQISIPYENCNDNPMYIC NKTKSCRSALDQNCQWEPNQEGLIALPENIGIWHLVGNLSKLTAKENYDNAL FCORNHALLASLTOKKVEFVFKOLRIMOSSQSMKLTLPWGLRKNVSYWCWDM SPETNSLIQWSEPSDAGFCGILSEBSTRGLKATCINPLNGSVCRERPAHSAKQC TPCLARTACGTGSSSCMCSNWKQCVDSNAYVSPFPQCMWNTMSTCPENCNS GYTCSCHEQFGCNCWDPSNTGKRCIEGYSKGPVMPQAPGTGNFYQPLNLSM CLEDSRYNWSFIHCPACNCHSKINQSIQCEKCNLTGKCECTCISGFYDPTNG KCQCKNCHASLNTNTGKCFCTTKYQKDECOLCEVKNRYQGNPLRGTCYITLLID YQFTSLSQEDDRYVTAINFVATPDQNRDLDFINASKNFNLNITWAASFSQTAG BEMPVVSXTNIKKEYDSFNEKDFRNHPNITFFVYVSNTFWPIKIQVTEQ"

## ORIGIN

Query Match 91.3%; Score 3487; DB 9; Length 3597;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 3524; Conservative 0; Mismatches 25; Indels 2; Gaps 2;

QY 270 GTCGGGCTCAGCGCGAGCGGCAAGAAATGTGACCGGCCCTGTGTCAACGGCGGTGCG 329  
 DB 48 GCGCGCGAGCGAGCGCTCGCGGCGAGGCGGCGCGGCGGCGCTGTGTCAACGGCGGTGCG 107  
 QY 330 CTGCAACCTGCGACCGGCGAGTGTGCTGTGCCCCCGCGGTGGGTGGCGGAGCAATGCCA 389  
 DB 108 CTGCAACCTGCGACCGGCGAGTGTGCTGTGCCCCCGCGGTGGGTGGCGGAGCAATGCCA 167

QY 390 GCACCTGCGGGGCGCGCTTCAGACTAACTGGATCTTCTGGTTTGTGACAGATGACCTGG 449  
 DB 168 GCACCTGCGGGGCGCGCTTCAGACTAACTGGATCTTCTGGTTTGTGACAGATGACCTGG 227  
 QY 450 AAATTATAATACAAACGAAGTGCAGTGGCTCATTTGAAGGACGACCAAAATAGATAAT 509  
 DB 228 AAATTATAATACAAACGAAGTGCAGTGGCTCATTTGAAGGACGACCAAAATAGATAAT 287  
 QY 510 GAGACTTCGTTTCAATCAATTTCTCAGAGTGTAGTTGGACCAATTAATATGTTATGA 569  
 DB 288 GAGACTTCGTTTCAATCAATTTCTCAGAGTGTAGTTGGACCAATTAATATGTTATGA 347  
 QY 570 TGGGACTCAATTTATGACCGCTAGTCTGCTCAATTTAGTGGCTCATTTGTTCTCTGAGAG 629  
 DB 348 TGGGACTCAATTTATGACCGCTAGTCTGCTCAATTTAGTGGCTCATTTGTTCTCTGAGAG 407  
 QY 630 AGATGGCAATGACAGCTGCTCCCTGAGGTGTGTCACATCAGGTTATGCCCTTGTGCAATTT 689  
 DB 408 AGATGGCAATGACAGCTGCTCCCTGAGGTGTGTCACATCAGGTTATGCCCTTGTGCAATTT 467  
 QY 690 TTTTAGTGATGCTGCTTATAATTTGACTGATTTAATATTACTTACAGTTTGTATGTG 749  
 DB 468 TTTTAGTGATGCTGCTTATAATTTGACTGATTTAATATTACTTACAGTTTGTATGTG 527  
 QY 750 TCCAAATAACTGCTCAGGCGGAGGAGTGTAAAGTCAAGTAAATAGCAGCAAACTGTGA 809  
 DB 528 TCCAAATAACTGCTCAGGCGGAGGAGTGTAAAGTCAAGTAAATAGCAGCAAACTGTGA 587  
 QY 810 ATGTGAATGTTTGAATACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869  
 DB 588 ATGTGAATGTTTGAATACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647  
 QY 870 CTGTGTTTTCCTCATCGAGGATCTGCAATTCAGTGTATGTCAGAGGATGCTCCCTGCTT 929  
 DB 648 CTGTGTTTTCCTCATCGAGGATCTGCAATTCAGTGTATGTCAGAGGATGCTCCCTGCTT 707  
 QY 930 CTCAGACTGCGGCGCTCTGGATGTTTCAGTTCCTGTACAGCTAACCACTCATTTTGGAC 989  
 DB 708 CTCAGACTGCGGCGCTCTGGATGTTTCAGTTCCTGTACAGCTAACCACTCATTTTGGAC 767  
 QY 990 TCAGAGGAATATTTAACTTAAAGTCCCGAGAGATCTATAAGCTGTGGTCAATGG 1049  
 DB 768 TCAGAGGAATATTTAACTTAAAGTCCCGAGAGATCTATAAGCTGTGGTCAATGG 827  
 QY 1050 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAAATGTTCT 1109  
 DB 828 AAACATTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAAATGTTCT 887  
 QY 1110 AGCGTATGACTTGTCTTAGGAGTGGCTTCCATTAACCGTTCTGTGAAACAATGTGT 1169  
 DB 888 AGCGTATGACTTGTCTTAGGAGTGGCTTCCATTAACCGTTCTGTGAAACAATGTGT 947  
 QY 1170 TGTTAGATATGTCATTTCTTTGTCATTATACAAGGATAAAATTTACATGTATGGAGAAA 1229  
 DB 948 TGTTAGATATGTCATTTCTTTGTCATTATACAAGGATAAAATTTACATGTATGGAGAAA 1007  
 QY 1230 AATTGATCAACTGGGAATGTGACCAATGAGTGTGAGGTTTTTTCATCAATTAATAGTGC 1289  
 DB 1008 AATTGATCAACTGGGAATGTGACCAATGAGTGTGAGGTTTTTTCATCAATTAATAGTGC 1067  
 QY 1290 ATGGGTGTTGTTGACCCCTTAAGGCAAGGAGAGTGTGAGTGGTGGGCACTCTGCACA 1349  
 DB 1068 ATGGGTGTTGTTGACCCCTTAAGGCAAGGAGAGTGTGAGTGGTGGGCACTCTGCACA 1127  
 QY 1350 CATTTGTACACTGAAGAATGGCGGAGTGTGCTGCTGGTTCATCTTTGGTCACTGCCCTCT 1409  
 DB 1128 CATTTGTACACTGAAGAATGGCGGAGTGTGCTGCTGGTTCATCTTTGGTCACTGCCCTCT 1187  
 QY 1410 CTATGATATATAGCAATGTGAGGAATATGATTTGGATAAGAACACATGAGTATATT 1469  
 DB 1188 CTATGATATATAGCAATGTGAGGAATATGATTTGGATAAGAACACATGAGTATATT 1247



QY 1470 ACACACCCAGGGTCCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTCTTACGACCATAG 1529  
Db 1248 ACACACCCAGGGTCCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTCTTACGACCATAG 1307  
QY 1530 GACCAGGCCCTTACGTTGATGTTGCTACAGGCTTTACGTCGCAATAGTACCGGCT 1589  
Db 1308 GACCAGGCCCTTACGTTGATGTTGCTACAGGCTTTACGTCGCAATAGTACCGGCT 1367  
QY 1590 TGCAGATGATCTCTACCGATATGTTGGATACCCAGATGTGGACCAATCTTAAAGACAG 1649  
Db 1368 TGCAGATGATCTCTACCGATATGTTGGATACCCAGATGTGGACCAATCTTAAAGACAG 1427  
QY 1650 CCGATTTTCCGTTACTTGGACACAGCTGTGATAGTGGAGAACCAATGCTGTGTTGG 1709  
Db 1428 CCGATTTTCCGTTACTTGGACACAGCTGTGATAGTGGAGAACCAATGCTGTGTTGG 1487  
QY 1710 GGGAAACACACAAATGACATCTATGAGCCATGCGGCCAAATGCTTCTTCCAGATTT 1769  
Db 1488 GGGAAACACACAAATGACATCTATGAGCCATGCGGCCAAATGCTTCTTCCAGATTT 1547  
QY 1770 CATGCCCTATGACATTTGCTGTGACCGCTGTGATGCTTCCAGACCTGATCTCCACCA 1829  
Db 1548 CATGCCCTATGACATTTGCTGTGACCGCTGTGATGCTTCCAGACCTGATCTCCACCA 1606  
QY 1830 TGATGTCAACAGATTTGSCCAAT-CAGCAGTCTTACACACAGCACCAGTATGTGTTGG 1888  
Db 1607 TGATGTCAACAGATTTGSCCAATCCAGCAGTCTTACACACAGCACCAGTATGTGTTGG 1666  
QY 1889 GTGGTTTCAATAGTCTCTCTCAGCGACATCTCTGGTATTCACCTCGGAAACAGTGTGATG 1948  
Db 1667 GTGGTTTCAATAGTCTCTCTCAGCGACATCTCTGGTATTCACCTCGGAAACAGTGTGATG 1726  
QY 1949 CGATTCGAGTGAAGCCGCTTGTAGCAGCAGACCTGTTATTCGGTGTGTTGGAAACA 2008  
Db 1727 CGCATTCGAGTGAAGCCGCTTGTAGCAGCAGACCTGTTATTCGGTGTGTTGGAAACA 1786  
QY 2009 CAGGTCCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAAGAAAGTTAA 2068  
Db 1787 CAGGTCCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAAGAAAGTTAA 1846  
QY 2069 AATCAGAAATGTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACAGCACAGATT 2128  
Db 1847 AATCAGAAATGTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACAGCACAGATT 1906  
QY 2129 GTTACAGCTGTACAGCCACCAACCAATGACTGCCATGGTGCATGACCAATGTCCTCCCA 2188  
Db 1907 GTTACAGCTGTACAGCCACCAACCAATGACTGCCATGGTGCATGACCAATGTCCTCCCA 1966  
QY 2189 GGAACCCAGCTGCTCAGAAGGCCAGATCTCCATTTTATGATGATGAGAAATGCCCCAAGG 2248  
Db 1967 GGAACCCAGCTGCTCAGAAGGCCAGATCTCCATTTTATGATGATGAGAAATGCCCCAAGG 2026  
QY 2249 ATAAACCCATGTAATCTGTAACAAGAACACAGCTGACAGAGCTGTGCCCTGACACAGA 2308  
Db 2027 ATAAACCCATGTAATCTGTAACAAGAACACAGCTGACAGAGCTGTGCCCTGACACAGA 2086  
QY 2309 ACTGCCAGTGGAGCCCGGAATCAGGATGCAATGCGCTGCCGAAATATCTGTGGCA 2368  
Db 2087 ACTGCCAGTGGAGCCCGGAATCAGGATGCAATGCGCTGCCGAAATATCTGTGGCA 2146  
QY 2369 TTGGCTGGCAATTTGGTTGGAATCTCATGTTTGAATAATTAATCTGCCAAGGAGAAATATG 2428  
Db 2147 TTGGCTGGCAATTTGGTTGGAATCTCATGTTTGAATAATTAATCTGCCAAGGAGAAATATG 2206  
QY 2429 ACAATGCTAAATGTTCTGTGGAACCAACAATGCCCTTTGGCTCTCTTACACCCAGA 2488  
Db 2207 ACAATGCTAAATGTTCTGTGGAACCAACAATGCCCTTTGGCTCTCTTACACCCAGA 2266  
QY 2489 AGAAGGTAGAAATTTGCTTTAAGCAGCTGCGAATAATGAGTCACTCTCAGAGCATGTCCA 2548  
Db 2267 AGAAGGTAGAAATTTGCTTTAAGCAGCTGCGAATAATGAGTCACTCTCAGAGCATGTCCA 2326  
QY 2549 AGCTACCTTAACCCCATGGGTGGCCCTTCGGAAAGATCAATGTGCTACTGCTGCTGGG 2608

Db 2327 AGCTCACCTTAACCCCATGGGTGGCCCTTCGGAAGATCAATGTGCTTACTGGTCTGGG 2386  
QY 2609 AAGATATGTCCCATTTTACAAATAGTTTACTACAGTGTGATGCGTCTGAGCCAGTGTG 2668  
Db 2387 AAGATATGTCCCATTTTACAAATAGTTTACTACAGTGTGATGCGTCTGAGCCAGTGTG 2446  
QY 2669 CTGGATTTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACTGCA 2728  
Db 2447 CTGGATTTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACTGCA 2506  
QY 2729 TCACACCCACTCAATGTAGTGTCTGTGAAGGCCCTGCAAAACCAACAGTGTAAAGCAGTGC 2788  
Db 2507 TCACACCCACTCAATGTAGTGTCTGTGAAGGCCCTGCAAAACCAACAGTGTAAAGCAGTGC 2566  
QY 2789 GGAACCACTGTGCTTTGAGGACAGCATGTGGAGATTGCAACAGCGGACAGCTCTGAGTGA 2848  
Db 2567 GGAACCACTGTGCTTTGAGGACAGCATGTGGAGATTGCAACAGCGGACAGCTCTGAGTGA 2626  
QY 2849 TGTGGTGCAGCACATGAGCAGTGTGTGACTCCAAATGCCCTATGTGGGCTCTCTTCCCTT 2908  
Db 2627 TGTGGTGCAGCACATGAGCAGTGTGTGACTCCAAATGCCCTATGTGGGCTCTCTTCCCTT 2686  
QY 2909 TTGGCCAGTGTATGGAAATGTTATACGATGAGCACCTGCCCCCTGAAAAATGTTCAAGGT 2968  
Db 2687 TTGGCCAGTGTATGGAAATGTTATACGATGAGCACCTGCCCCCTGAAAAATGTTCAAGGT 2746  
QY 2969 ACTGTACTGTAGTCAATTCCTTTGGACACAGGCTGTGGTGTGATGTGATCCACAGA 3028  
Db 2747 ACTGTACTGTAGTCAATTCCTTTGGACACAGGCTGTGGTGTGATGTGATCCACAGA 2806  
QY 3029 ATACTGGCAAGGAAATGCATAGAGGTTTCTTAAAGGACCACTGAAAGTGCCTTCGC 3088  
Db 2807 ATACTGGCAAGGAAATGCATAGAGGTTTCTTAAAGGACCACTGAAAGTGCCTTCGC 2866  
QY 3089 AAGCCCTCACAAGAAATTTCTATCCACAGCCCTGCTCAATTCAGCATGTGTCTAGAG 3148  
Db 2867 AAGCCCTCACAAGAAATTTCTATCCACAGCCCTGCTCAATTCAGCATGTGTCTAGAG 2926  
QY 3149 ACAGCAGATACACTGCTCTTTCATTCAGTCTCCAGCTTGCCTAAATGCAAGGCAACAGTA 3208  
Db 2927 ACAGCAGATACACTGCTCTTTCATTCAGTCTCCAGCTTGCCTAAATGCAAGGCAACAGTA 2986  
QY 3209 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCAACAGGCAAGCACTGG 3268  
Db 2987 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCAACAGGCAAGCACTGG 3046  
QY 3269 AGACTGTCAATCTGGCTTCTACGTTGATCCCAACATGGAGGAAATGTCAGCCATGCA 3328  
Db 3047 AGACTGTCAATCTGGCTTCTACGTTGATCCCAACATGGAGGAAATGTCAGCCATGCA 3106  
QY 3329 AGTGCATGCGCACGCTCTCTGTGCAACACCAACACGCGGCAAGTGTCTTCGACCAACCA 3388  
Db 3107 AGTGCATGCGCACGCTCTCTGTGCAACACCAACACGCGGCAAGTGTCTTCGACCAACCA 3166  
QY 3389 AGGCGCTCAAGGGGACGAGTGCAGTATGTGAGGTAGAAATTCGATACCAAGGAAACC 3448  
Db 3167 AGGCGCTCAAGGGGACGAGTGCAGTATGTGAGGTAGAAATTCGATACCAAGGAAACC 3226  
QY 3449 CTCTCAGAGAAATGTTATATCTTCTTATGACTATCAGTTTCACTTCACTTGTCTAT 3508  
Db 3227 CTCTCAGAGAAATGTTATATCTTCTTATGACTATCAGTTTCACTTGTCTAT 3286  
QY 3509 CCCAGGAAAGTATGCTGATTAACAGCTATCAATTTTGGTATCTCTGAGCAACAAA 3568  
Db 3287 CCCAGGAAAGTATGCTGATTAACAGCTATCAATTTTGGTATCTCTGAGCAACAAA 3346  
QY 3569 ACAGGATTTGGACATGTTTATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGG 3628  
Db 3347 ACAGGATTTGGACATGTTTATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGG 3406  
QY 3629 CTGCACTGCTCAGCTGGAAACCCAGGCTGGAGAGATGCTCTGTTGTTTCAAAAACCA 3688

```

Db      3407 CTGCCAGTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCA 3466
QY      3689 ACATTAAGAGATCAAAAGATAGTTTCTTAATGAGAAGTTGATTTTCGCAACCAACCCAA 3748
Db      3467 ACATTAAGAGATCAAAAGATAGTTTCTTAATGAGAAGTTGATTTTCGCAACCAACCCAA 3526
QY      3749 ATATCACTTTCTTTGTTATGTCAGTAATTTTCACTTGGCCCAATCAAAATTCAGGTGCAAA 3808
Db      3527 ATATCACTTTCTTTGTTATGTCAGTAATTTTCACTTGGCCCAATCAAAATTCAGGTGCAAA 3586
QY      3809 CTGAACAATGA 3819
Db      3587 CTGAACAATGA 3597

RESULT 5
AF531101
LOCUS      4493 bp mRNA linear MAM 19-AUG-2002
DEFINITION Bos taurus attractin mRNA, complete cds.
ACCESSION AF531101
VERSION AF531101.1 GI:22297307
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 4493)
REFERENCE
  Graphodatskaya,D., Joerg,H. and Stranzinger,G.
  Direct Submission
  Submitted (18-JUL-2002) Animal Sciences, ETH, Tannenstrasse 1,
  Zurich, CH 8092, Switzerland
  Location/Qualifiers
    1. 4493
      /organism="Bos taurus"
      /mol_type="mRNA"
      /db_xref="taxon:9913"
      /chromosome="13"
      52..4299
        /note="transmembrane form; mahogany protein"
        /codon_start=1
        /product="attractin"
        /protein_id="AA055445.1"
        /db_xref="GI:22297308"
        /translation="MVAATEARLRIVTVTPAPTRNRGRRCRDWDATRAGRPLRLTGL
        RLPRWRRLILLPLLLLPAAEAAAASVSGSAAEAKEDRCPVNGGRCPGT
        GQVCPAGWGEQCHQGRFLTGSGFITDGPNGYKTKCTWLIEGPNRIMRLR
        FNHPATECSWHLVYDGDLSIYAPLVAFAFSLIYPERDGNETVPEVVAATSGYALLHFF
        SDAYNLGTNITNFDMCPNCSGRGCEKLSNSNTVQCECSNWKEACDIPHCYN
        NCGPFRIGCINSSDRCSEWQPGCVFPANQSFWRTEYSIDLKPLRASHKAV
        VNGIMVVGYNHSDYNMVLADLASREWALNRSVNSVVRVYHGSLLALYODKIY
        MYGKIDSTGNVRELRFVHINSEWLLSPKAEQYAVVGHSAHIVTLKGRVVMVL
        IFGHCPLYGYSVQVEYDLDRNTWSILHGTALVQGGYHGSVVDRLKALYIHGGYK
        AFSANKYRLADLVYDVTOMWTILKDSRFRVLTAVIVSGTLMVFGGNTNDTSM
        SHGAKFSSDFMAYDIDCRWSVLPRLDHDVNRFGHSVAVLNRYVFGFNSLL
        SDILVFTSECEAEQSENAACLAAGPVRCDMDTSSQCVSEWELAPAEQKIKSECRSK
        RIFDHRCDQDITICNTANTGCMNDHCVPNHSCTEQGISIFKIDPCNDMPY
        YCNKTSRCSALDONQWEPNRCIALPENICGIMHVLVGNLSCLITTARTEDNA
        KLSRNHNALASLTITQKKEVFLKQRLIMQSSQMSKFTLTPWGLRLKLNVSYWWE
        DMSPTFSLQWMESEPSDAGFCGILSEPSRGLKAATCINPLNGSVCEPANHSAKQ
        CRTPCALRTACGETSGSCECMWCKQCVDSNAYVSPFGOCMEWYMSVCPNEN
        CSYCTSHCLEQPCQWCTDPSNTGKGIKESYKGVKMPGQSGPTGNSVPOPLNS
        SMCLDSRYNWSFTHCACQNGHSHKINSQICEKCNLTIVGHKCEICISGFQDPTN
        GKGQPCRCNGHSLCNTNGKCTCTTGNGVDEQCEVENRYQGNPLAGTCTYILL
        IDYQFTSLSQEDDRIYTAINFATPDQRNLDLMTFNASKNFMNLTWAFSFSAGTF
        AGEEMPVVSKTNIKEYKDSFSENFDPNRHFNITFFVYNSNFTWPKIQIAPFSQNF
        MDLVQFFVTFPSFCLSLILVAAVVKIKQSCWASRRREQLIREMQQWASRPFASVNA
        LETDEPDLITGSGIKTVPKPIALEPCFGNKAALVSFVRLPRGLGGIPPGQSGLA
        ASALVDISQMPVVVYKESGAVNRNRKQPPAQPTCI"
  ORIGIN
    85.1%; Score 3251.8; DB 4; Length 4493;
    Best Local Similarity 92.0%; Pred. No. 0;

```

```

Matches 3488; Conservative 0; Mismatches 267; Indels 36; Gaps 4;
QY      13 GCAGCGGCAACAGGCTGAGAGAGAGACCGCGCGAGCGGACGCTCGCGGGC 72
Db      58 GCAGCGGCAACAGGCTGAGAGAGAGACCGCGCGAGCGGACGCTCGCGGGC 114
QY      73 AGAGAGCGGCGCGGCGGCTGAGAGAGAGACCGCGGCTGGAGAGCGGCGGCTGGGG 132
Db      115 AGGAACGCGAGCGGCTGAGAGAGAGACCGCGGCTGGAGAGCGGCGGCTGGGG 174
QY      133 GCGGCGGCTGCGCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCGGCTGCTGCTG 192
Db      175 ACCGGGCTGCGCTCCCGCGGCTGCTGTCTCCACCGCTGCGGCGGCTGCTGCTG 213
QY      193 CTGTTGTTGCTCCCGCGGCGGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 252
Db      214 CTGCTGCTGCTCCCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 264
QY      253 GCGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 312
Db      265 GCGGCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 324
QY      313 TGTGTCAACGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372
Db      325 TGTGTCAACGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 384
QY      373 GTCGGCGAGCAATGCCAGCACTGCGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 432
Db      385 GTGGGCGAGCAATGCCAGCACTGCGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTG 444
QY      433 GTGACAGATGGACCTGGAATATTAATAACAAAGCAAGTGCACGCTGCTGCTGCT 492
Db      445 ATAACAGATGGACCTGGAATATTAATAACAAAGCAAGTGCACGCTGCTGCTGCTG 504
QY      493 CAGCCAAATAGATAATAGACTTGTGTTCAATCAATTTTGTCTACAGAGTGTAGTTGG 552
Db      505 CAGCCAAATAGATAATAGACTTGTGTTCAATCAATTTTGTCTACAGAGTGTAGTTGG 564
QY      553 CATTTATATGTTTATGATGGGACCTCAATTTATGACCGCTAGTTGCTCATTAGTGG 612
Db      565 CATTTATATGTTTATGACGGGACCTCAATTTATGACCGCTAGTTGCTCATTAGTGG 624
QY      613 CTCATTGTTCTGAGAGAGATGGCAATAGACTGTCCTGAGAGTGTGTCACATCAGT 672
Db      625 CTCATTGTTCTGAGAGAGATGGCAATAGACTGTCCTGAGAGTGTGTCACATCAGT 684
QY      673 TATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
Db      685 TATGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 744
QY      733 TACAGTTTGTATGTTGCTCAATAATGCTGCTCAGCGGAGAGAGTGAAGATCAGT 792
Db      745 TACAATTTTGACATGTTGCTCAATAATGCTGCTCAGCGGAGAGATGAAGATCAGT 804
QY      793 ACAGCGAACTGTTCAATGAAATGTTTGAATACTGGAAGAGTGAAGCATGTCATT 852
Db      805 AGAGCAACACTGTTCAATGAAATGTTTGAATACTGGAAGAGTGAAGCATGTCATT 864
QY      853 CCTCACTGTACAGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912
Db      865 CCTCACTGTGTAATAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 924
QY      913 AGAGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 972
Db      925 AGAGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 984
QY      973 AACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1032
Db      985 AACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044
QY      1033 AAAGCTGTGGTCAATGGAAACATTAATGTTGGGTTGTTGGAGGATATATTTCAAC 1092
Db      1045 AAAGCTGTGGTCAATGGAAACATTAATGTTGGGTTGTTGGAGGATATATTTCAAC 1104

```

1093 GATTATACATGGTTCTAGACGATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCGT 1152  
1105 GATTACAACTGGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGCTTGCACGACCGT 1164  
1153 TCTGTGAACAATGGTTGTTAGATATGCTCATTTCTTTGGCATATATACAGGATAAAATT 1212  
1165 TCTGTGAACAATGGTTGTTAGATATGCTCATTTCTTTAGCATATATACAGGATAAAATT 1224  
1213 TACATGTATGAGGAGAAATTTGATCCAACTGGGAATGACCAATGATGTAGAGTGTGTTT 1272  
1225 TACATGTATGAGGAGAAATTTGATCCAACTGGGAATGACCAATGATGTAGAGTGTGTTT 1284  
1273 CACATTCATATGATGATGGTGTGTTGACCTTAGGCAAGGAGGACGATGACAGTG 1332  
1285 CATATTCATATGATGATGGTGTGTTGACCTTAGGCAAGGAGGACGATGACAGTG 1344  
1333 GTTGGGCACTCTGCACACATTTGTTACACTGAAGATGGCCGAGTGGTCACTGCTGTCATC 1392  
1345 GTTGGGCACTCGGCGCACATTTGTCACGTTGAAATGGCCGAGTGGTCACTGCTGTTATC 1404  
1393 TTTGGTCACTGCTCTCTATGGATATATATAGCAATGTGCAAGATATGATTTGGATAAG 1452  
1405 TTTGGTCACTGCTCTCTATGGATATATATAGCAATGTGCAAGATATGATTTGGATAAG 1464  
1453 AACACATGGAGTATATACACACCGAGGTGCTTGTGCAAGGGGGTTACGGCCATAGC 1512  
1465 AACACATGGAGTATATACACACTCAGGGTGCTTGTTCRAAGGGGCTACGGCCATAGC 1524  
1513 AGTGTATACGACCATAGGACCGCCCTATACGTTTCATGTTGCTGCTCAAGGCTTTGAGT 1572  
1525 AGTGTATGATGATAGGACCAAGCCCTGTATATTCACGGAGGCTACAAAGGCTTTGAGT 1584  
1573 GCCAATAGTACGGCTGCAATGATCTCTACCGATATGATGATGATGATGATGATGATGATG 1632  
1585 GCCAATAAATACCGGCTGCAATGATCTCTACAGATATGATGATGATGATGATGATGATGATG 1644  
1633 ACCATTCCTTAAGGACAGCGATTTTTCGGTTACTTTCGACACAGCTGTGATGATGATGATG 1692  
1645 ACCATTCCTTAAGGACAGCGATTTTTCGGTTACTTTCGACACAGCTGTGATGATGATGATG 1704  
1693 ACCATGCTGGTGTGTTGGGGAAACACACACATGACATCTATGAGCCATGGGCCAAA 1752  
1705 ACCATGCTGGTGTGTTGGGGAAACACACACATGACATCTATGAGCCATGGGGCCAAA 1764  
1753 TGTCTCTTCAGATTTTCATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1812  
1765 TGTCTCTTCAGATTTTCATGCTTATGATGATGATGATGATGATGATGATGATGATGATGATG 1824  
1813 AGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGAGTCTTACACACAGC 1872  
1825 AGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGAGTCTTACACACAGC 1884  
1873 ACCATGATGTTGCTGGTGGTTTCAATAGTCTCTCCCTCAGGACATCTGTTATTCACC 1932  
1885 ACCATGATGTTTGGTGGTGGTTTCAATAGTCTCTCCCTCAGGACATCTGTTATTCACA 1944  
1933 TCGGAACAGTGTGATCGCATCGGAGTGAAGCCGCTGTTTATGACGAGGACCTGGTATT 1992  
1945 TCGGAACAGTGTGATCGCATCGGAGTGAAGCTGTTGTTCTCGGAGCAGGACCTGGTATT 2004  
1993 CGGTGTGTGGAACACAGGGTGTCTCAGTGTATCTGTTGGGCGCTGGCACTGATGAA 2052  
2005 CGGTGTGTGGAACACAGGGTGTCTCAGTGTATCTGTTGGGCGCTGGCACTGATGAA 2064  
2053 CAAGAAGAAAGTTAAATCAGATGTTTTCCTCAAGAACTCTTGACCATGACAGATGT 2112  
2065 C---AAGAAGAAATGAAGTCAAGATGTTTCTTAAGAAATTTTGACCATGACAGATGT 2121  
2113 GACCAACACAGATTTGTACAGCTGTACAGGCAACACCAATGACCTGCCACTGGTGCAAT 2172  
2122 GACCACTACAGGATTTGTACAGCTGTACAGGCAACACCAATGACCTGCCACTGGTGCAAT 2181

2173 GACCATTTGTCCTCCAGGAACACACAGCTGCTCAGAAAGGCGAGATCTCCATTTTAGGTAT 2232  
2182 GACCATTTGTCCTCCAGGAACACACAGCTGCTCAGAAAGGCGAGATCTCCATTTTAGGTAT 2241  
2233 GAGAATTTGCCCAAGGATAACCCCATGTACTACTGTAAACAAGAACCCAGCTGCGAGGAC 2292  
2242 GATCATTTGCCCAAGGATAACCCCATGTATTTATTGTAAACAAGAACCCAGCTGCGAGGAC 2301  
2293 TGTGCCCTCGGACCAAGCTGCCAGTGGGAGCCCGGAATCAGAGTGCATTTGCCCTGCC 2352  
2302 TGTGCCCTCGGACCAAGCTGCCAGTGGGAGCCCGGAATCAGAGTGCATTTGCCCTGCC 2361  
2353 GAAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACTCATGTTTGAATTTACTACT 2412  
2362 GAAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACTCATGTTTGAATTTACTACT 2421  
2413 GCCAAGGAGATTTATGACATGCTAAATTTGTTCTGTAGGAACCAACATGCCCTTTGGCT 2472  
2422 GCCAAGGAGATTTATGACATGCTAAATTTGTTCTGTAGGAACCAACATGCCCTTTGGCT 2481  
2473 TCTCTTACAAACCCAGGAAGAGGTAGAAATTTGCTTAAAGCAGCTGCGAAATTCAGTCA 2532  
2482 TCTCTTACAACTCAGAAAGGTAGAAATTTGCTTAAAGCAGCTGCGAAATTCAGTCA 2541  
2533 TCTCAGAGATGTCGAAGCTCACTTTAAACCCATGGGTGGCTTCGGAAGATCAATG 2592  
2542 TCAAGAGATGTCGAAGCTCACTTTAAACCCATGGGTGGCTTCGGAAGATCAATG 2601  
2593 TCTCTGCTGGTGGGAGGATATGTCCTTAAACAAATTTACTTACAGTGGATGCCG 2652  
2602 TCTCTGCTGGTGGGAGGATATGTCCTTAAACAAATTTACTTACAGTGGATGCCG 2661  
2653 TCTGAGCCAGTGTGATGCTGGAATTTTATCGAAACCCAGTACTCGGGAGCTG 2712  
2662 TCTGAGCCAGTGTGATGCTGGAATTTTATCGAAACCCAGTACTCGGGAGCTG 2721  
2713 AAGCTCTCAACTGCTCAACCCAGTACTGCTGTAAGAGGCTGCAAAACCC 2772  
2722 AAGCTCTCAACTGCTCAACCCAGTACTGCTGTAAGAGGCTGCTAACCC 2781  
2773 AGTCTAAGCAGTGGCCGACACCATGCTTGGAGCAGCAGTGTGGAGATTCACACAGC 2832  
2782 AGTCTAAGCAGTGGCCGACACCATGCTTGGAGCAGCAGTGTGGAGATTCACACAGT 2841  
2833 GGCAGCTCTGAGTGTGCTGGTGGAGCAACATGAGCAGTGTGAGCTTCAATGCTAT 2892  
2842 GGCAGCTCTGAGTGTGCTGGTGGAGCAACATGAGCAGTGTGAGCTTCAATGCTAT 2901  
2893 GTGGCTCTCTTCCCTTTGGCCAGTGTATGGAATGTTATACGATGAGCAGCTGCCCCCT 2952  
2902 GTGGCTCTCTTCCCTTTGGCCAGTGTATGGAATGTTATACGATGAGCAGCTGCCCCCT 2961  
2953 GAAAAATTTTTCAGGCTACTGTACTGTAGTCAATTTGTTGGAGCAACAGGCTGTGGCTGG 3012  
2962 GAAAAATTTTTCAGGCTACTGTACTGTAGTCAATTTGTTGGAGCAACAGGCTGTGGCTGG 3021  
3013 TGTACTGATCCAGCATACTGCAAGGGAATGATAGAGGTTCTCTATAAGAGGACCA 3072  
3022 TGTACTGATCCAGCATACTGCAAGGGAATGATAGAGGTTCTCTATAAGAGGACCA 3081  
3073 GTGAAGATGCTTCCGAGCCCTTACAGGAAATTTCTATCCAGCCCTGCTCAATTC 3132  
3082 GTGAAGATGCTTCCGAGCCCTTACAGGAAATTTCTATCCAGCCCTGCTCAATTC 3141  
3133 AGCATGTGTAGAGGACAGCAGATACAACTGGTCTTTCACTTCACTTCCAGCTTGCACA 3192  
3142 AGCATGTGTAGAGGACAGCAGATACAACTGGTCTTTCACTTCACTTCCAGCTTGCACA 3201  
3193 TGCACGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGAC 3252  
3202 TGCATGCGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGAC 3261  
3253 ACAGCAAGCACTCGGAGACCTGCTATCTGGCTTCTACGGTGTATCCCAATGGAGG 3312

```

Db      3262  ACGGGAAGCACTGTGAGACCTGCATATCTGGCTTATGCTGATCAACCAACGGGGA 3321
QY      3313  AAATGTCAGCATGCAAGTGCAGTGGACGGCTCTCTGTGCAACACCAACGGGCAAG 3372
Db      3322  AAATGTCAGCATGCAAGTGCAGTGGACGGCTCTCTGTGCAACACCAACGGGCAAG 3381
QY      3373  TGCTTCTGCAACCAAGGCGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAT 3432
Db      3382  TGCTTCTGCAACCAAGGCGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAT 3441
QY      3433  CGATACCAAGGAAACCTCTCAGAGGAAATGTTATTAATCTTCTTATGACTATCAG 3492
Db      3442  CGATACCAAGGAAACCTCTCAGAGGAAATGTTATTAATCTTCTTATGACTATCAG 3501
QY      3493  TTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGGCT 3552
Db      3502  TTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGGCT 3561
QY      3553  ACTCTGAGCAACAAACAGGATTTGAGATGTTTATCAATGCTCTCAAGAAATTCAC 3612
Db      3562  ACACCTGATGAACAAACAGGATTTGAGATGTTTATCAATGCTCTCAAGAAATTCAC 3621
QY      3613  CTCACATCACTGGGCTGCCAGTTTCTCAGCTGGAACCCAGCTGGAGAGAGATGCT 3672
Db      3622  CTCACATCACTGGGCTGCCAGTTTCTCAGCTGGAACCCAGCTGGAGAGAGATGCT 3681
QY      3673  GTTGTTCACAAACCAACATTAAGAGTACAAGATAGTTCTTCTPAATGAGAAGTTGAT 3732
Db      3682  GTTGTTCACAAACCAACATTAAGAGTACAAGATAGTTCTTCTPAATGAGAAGTTGAT 3741
QY      3733  TTTGCAACACCCCAATATCACTTCTTGTGTTATGATGAGTAATTCACCTGGCCATC 3792
Db      3742  TTTGCAACACCCCAATATCACTTCTTGTGTTATGATGAGTAATTCACCTGGCCATC 3801
QY      3793  AAAATTCAGT 3803
Db      3802  AAAATTCAGT 3812

```

```

RESULT 6
AF116897
LOCUS      6083 bp mRNA linear ROD 20-MAR-1999
DEFINITION Mus musculus mahogany protein mRNA, complete cds.
ACCESSION AF116897
VERSION AF116897.1 GI:4454560
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 6083)
AUTHORS Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr.,
DiRocco,L., Holmgren,L., Montegno,J., Bork,P., Huszar,D.,
Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L.,
Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.
TITLE The mahogany protein is a receptor involved in suppression of
obesity
JOURNAL Nature 398 (6723), 148-152 (1999)
MEDLINE 99184159
PUBMED 10086355
REFERENCE 2 (bases 1 to 6083)
AUTHORS Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr.,
DiRocco,L., Holmgren,L., Montegno,J., Bork,P., Huszar,D.,
Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L.,
Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.
TITLE Direct Submission
JOURNAL Submitted (28-DEC-1998) Sequence Analysis, Millennium
Pharmaceuticals, 640 Memorial Drive, Cambridge, MA 02139, USA
FEATURES
source
/organism="Mus musculus"
/mol_type="mRNA"

```

## CDS

```

/db_xref="taxon:10090"
/notes="MG"
/codon_start=1
/product="mahogany protein"
/protein_id="AAD20947.1"
/db_xref="GI:4454561"
/translation="MVAVAAATAAARLRGSIITTAAPAGRKGRQHRPCTATGAWRPG
PRARLCPLRVLSRALPPPLPLLSLILPLPREAAAVAAVAASAAAABAKCQR
PCVNGCRNPGTCQVCPTGCGEQCHGGRFLTNGSGSFVDTGPNKYTKCTMWL
IEQPNRIMRLRENHATPCSMDHLVYDGDSDIYAPLIAAFSLGVIPERDGNETAPEV
TVTSGTALHFFSDAAYNLTGFTITTFDMCNNSGRGECSSAVECECEENW
KGSQSDIPIHCTDNGCFPHRGICNASDTRGSCFPHWGPCSIPVPANOSFVTRBYS
DLKLPRASHKAVVNGNIMVVGGMFNHSDYSMLAYDILTSEMLPNHNSVNSVVRVY
GHSIALHKDKIYMGKGIISTGNTVNEFLVHINESWLLTPKAKDQAVVYGHSAH
VTLASGRVVMVLVI FGHCPYGVISVQVEYDLEKNTWSILHTOGALVQGGYGHSSVYDD
RTKALYVHGKYKAFSANKYRLADDLYEYVDVTQMWTLLKDSRFRVLTAVTAVSTML
VFGGNTHTNSHGAKCFSSDPMAYDIACDRMSVLPRLHHDVNRFGHSAVLNST
MYVFGFNSLLLSVLTSEQCDARHSEAAACVAAGFGRKLDWTSQTSRCSWEIATE
EQAEKLKSCFSTRKTLDDHDCDQHTDCYCTANTNDCHNDHCVPVNSHTEGQISI
AKYESCPKDPYCNKTSRSCALDQCNQWEPNQCETALPENICMGMHLVGNSC
LKITTAKENYDNALSCRHNHPLASLTQKVEFVLKQRLMQSQQSKLTLTPWV
GLRKINVSVMCEDMSPTNSLLOWPSPSDAGFCGILSEPTSGLKAATCINPLNG
SVCEPANSAKOCRTPCALRTACGCTSSSECMWCMKOCVDSNVAVSPFGQ
MEWYTSXCPPEPCSYCTCSHLEPGGQWCTDPSNTGKCIKESYKGPVMPBQA
SAGNVPQPLNNSMCLDSRYNSFIHPACQNGHSCINQSCICEKEDCTTGKHC
ETCISGFDPTNGGKQCPKNGHASLCTNTGKCTTGKVGDEDCQCEVENRYQ
GNPLGCTCYTLLIDYQTFSLQEDDRYTAINFVATPEQNRDLDMFINAKNFKL
NITWATSPAGTGTGEEVPSKTKIKEYDSFNEKDFRNFHNPITFFVYSNFTWP
KIITAFSQHNSFMDLVQFVTFKSCFLSLVAAVWKIKOSCWASRRRQLREMO
QNASRFSANVALETDEBPDLIGSIIKTVKPIALEPCFNGKAAVLSVFLPRGL
GATPPGQSRSCGCGPGHFPADANSVQGVRSCKKPEAAACTAWNLHLILGPEFA
LTGFWSTKLTSTGEARAKQAMRRRLGTLEASNSCA"

```

## ORIGIN

```

Query Match      82.4%; Score 3147.4; DB 10; Length 6083;
Best Local Similarity 89.5%; Pred. No. 0;
Matches 3401; Conservative 0; Mismatches 391; Indels 9; Gaps 1;

QY      3  GGTGGCGGAGCGCGGCAACTGAGCAAGGCTGAGGAGAGAGACGGCGCGACGCGCAGC 62
Db      78  GGTGGCGGCGCGCGCGGCGGCGGCGGCGGCTGAGGGGGAGACGACACGACGACGACG 137
QY      63  GCTCGCGGCGAGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 122
Db      138  GCCTGCGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 197
QY      123  GGGGCTGGGGGCGGGGCTCGGCTCCCGCGGCTGCTGTCTCCACCGCTGCGCGCCACGGCT 182
Db      198  GGGACCGCGCGCGCGGCTGTCTCCCGCGGCTGTCTCGCGGGGCTGTCTCGCGGGCGCT 257
QY      183  GCTGCTGCTGCTGTGTTGCTCCCGCGCGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 242
Db      258  GCTGCTGCGCGCTGCTCTT-----TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 308
QY      243  CGAGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 302
Db      309  CGAGGCGCGCTGCGGTGGCGGCGGCGGCTGTCCGCTCGGCGCGGCGGCGGCGGCGGCG 368
QY      303  TGACCGGCGCTGTGTCAACGCGGCTGTGTGCAACCGCTGSCACCGGCGAGTGTGTGCCC 362
Db      369  TGACCGGCGCTGTGTCAACGCGGCGGCTGTGTGCAACCGCTGSCACCGGCGAGTGTGTG 428
QY      363  CGCGGCGCTGGTGGGCGAGCAATGCGACACCTGCGGGGCGCGCTTTCAGACTAACTGGATC 422
Db      429  CAGGCGCTGGTGGGCGAGCAATGCGACACCTGCGGGGCGCGCTTTCAGACTAACTGGATC 488
QY      423  TTCTGGGTTTGTACAGATGAGACCTGGAAATTTATTAATACAAAAGAGTGCACGTGGCT 482
Db      489  TTCTGGATTTGTAAACAGATGAGACCTGGAAATTTATTAATATAAGACAAAGTGCACAT 548
QY      483  CATTGAAGGACACCAATAGAAATATGAGACTTCTGTTTCAATCATTTTGTGTACAGATG 542

```

549	Db	CATTGAAGCAGACCCAAATAGAAATATGAGACTTCGCTTCAACCAATTTTTGCTACAGAAATG	608
543	Qy	TAGTTCGGGACCAATTTATATATGTTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC	602
609	Db	TAGCTGGGACCAATTTATATATGTTTTATGATGGGACTCAATCTACGCACCTCTGATTTGCTGC	668
603	Qy	ATTTAGTGGCCTCATTTGTTCTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGTTGTTGC	662
669	Db	CTTTAGTGGCCTCATTTGTTCTCTGAAAGAGATGGCAATGAGACGGCTCTGAGGTCACATGT	728
663	Qy	CACATCAGGTTATGCTTGCTGCAATTTTTTTTAGTGATGCTGCTTATAATTTGACTGGATT	722
729	Db	CACCTCAGGTTATGCACTGCTGCAATTTTTTTTCAGTGATGCTGCTTATAATCTGACTGGATT	788
723	Qy	TAATATTACTTACAGTTTGTGATATGTTGCTCAATAATACTCTCAGGCCGAGGAGAGTGA	782
789	Db	TAATATCACTTACAATTTTGACATGTGTCCGAAATATTTGCTCAGGCCGAGGAGAGTGA	848
783	Qy	GATCAGTAAATAGCAGCGAAACTGCTGAATGTGAATGTTCTGAAAACTGGAAGGTGAAGC	842
849	Db	GAGCAGTTAACAGCAGACGCGCTGTTGAGTGTGAATGTTCTGAAACTGGAAAGGGAGTTC	908
843	Qy	ATGTGACATTCCTCACTGTACAGACCAACTGTGGTTTTCTCTCATCGAGGCAATCTGCAATTC	902
909	Db	GTGTGACATTCCTCACTGTACAGACCAACTGTGGCTTTCTCTCACCAGGAGCATCTGAATGC	968
903	Qy	AAGTGATGTACAGAGGATGCTCTGCTTCTCAGACTGCGAGGGTCTGTGATGTTCAAGTTC	962
969	Db	AAGCGATACAGAGGGTGTCTCTGCTTCTCTACTGGCAGGGTCTGTGATGTTCAATTC	1028
963	Qy	TGTACCAGCTAAACCAAGTCATTTTTGGACTCGAGAGGAATATTTAACTTAAAGCTCCCCAG	1022
1029	Db	TGTGCCAGCTAAACCAAGTCATTTTTGGACTCGAGAGGAATATTTCTGATTTAAAGCTTCCCAG	1088
1023	Qy	AGCATCTCTAAAGCTGTGTGTCATGGAAACATATATGTGGTGTGTTGGAGGATATATGTT	1082
1089	Db	AGCCTCTCTAAAGCTGTGTGTCATGGAAATATAATGTGGTGTGTTGGCGGATATATGTT	1148
1083	Qy	CAACCACTCAGATTTATAACATGGTTCTAGCGTATGACCTTGCTTCTTAGGAGTGGCTTCC	1142
1149	Db	CAACCATTCAGATTTACAGCATGGTTTATAGCTATGACCTGACCTCTTAGGGAATGGCTTCC	1208
1143	Qy	ACTAAACGGTTCGTGAAACAAATGTGGTGTAGATATGTCATCTTTGGCATTATACAA	1202
1209	Db	ACTAAACCATTCGTGAAACAGTGTGGTGTGAAAGATATGTCATCTTTGGCATTACATAA	1268
1203	Qy	GGATAAAATTTACATGTATGAGGAAAAATTGATCCAACTGGGAATGTGACCAATCAGTT	1262
1269	Db	GGATATAAATCTACATGTATGGAGAAAAATTTGATTCACAGGGAACTGACCAATCAGCT	1328
1263	Qy	GAGAGTTTTTCAATTCATAATAGTCAATGGGTGTGTTGACCCCTTAAGCCAAAGGAGCA	1322
1329	Db	GAGAGTATTTCAATTCATAATAGTCAATGGGTATTTGTTAACTCCGAAAGCTAAGGATCA	1388
1323	Qy	GTATGCACTGTGGGCACTCTGCACACATTTGTATACACTGAAGAAATGGCCGAGTGTGTCAT	1382
1389	Db	GTATGCACTGTGGTGGCACTCAGCACACATTTGTATACCTGGCATCTGGCCGTGTGTGTCAT	1448
1383	Qy	GCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGAGGAATATGA	1442
1449	Db	GTGTGTCACTTTGGTCACTTTGCCCACTCTATGGATATAAAGCGTTGTGAGGAATATGA	1508
1443	Qy	TTTGGATTAAGAACACATGGAGTATATTACACACCGAGGTGCCCTTGTGCAAGGGGGTTA	1502
1509	Db	CTTGGAAAGAACACATGGAGTATATTACATACTCAGGGTGTCTCTTTGTGCAAGGGGGTTA	1568
1503	Qy	CGGCCATAGCAGTGTTTACGACCATAGGACCCAGGGCCCTATAGCTTCAATGGTGGCTACAA	1562
1569	Db	TGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCAATGGTGGCTACAA	1628
1563	Qy	GGCTTTCAGTGCCAATAATAGTACCGGCTTGAGATGATCTCTACCGATATGATGTGGATAC	1622
1629	Db	GGCTTTCAAGCGCAACAATAACCGGCTTGAGATGAGCTCTCTACAGATACAGATGGGATAC	1688

Qy	1623	CGAGTGTGGACCAATCTCTTAAAGACAGCCGATTTTTCCGTTACTTTGCGACACACAGCTGTGAT	1680
Db	1689	TCAGATGTGGACCAATCTCTTAAAGACAGCCGATTTTTCCGTTACTTTCCGTTACTTTGCGATACAGCTGTGAT	1748
Qy	1683	AGTGAGTGGAAACCATCTCGTGTGTTGGGGGAAACACACAATAGCACATCTATGAGCCA	1742
Db	1749	AGTGAGTGGAAACCAATCTCGTGTGTTGGAGGGAACACACAATGACACATCTCCATGAGCCA	1808
Qy	1743	TGCGGCCAAATGCTTCTCTTCAGATTTCATTTGGCGTATGACATTTGCCCTGTGACCGCTGGTCT	1802
Db	1809	CGGTGCCAAATGCTTCTCTCGGACATTCATGGGCTTATGACATTTGTCATTTGTCGCGATGGTCT	1868
Qy	1803	AGTGCTTCCAGACCTGATCTCCACCATGATGTCACAGATTGGCCATTTACGACGTCTTT	1862
Db	1869	AGTGCTTCCAGACCTGAGCTCCATCATGATGTCAACAGATTGGCCATTTACGACGTCTTT	1928
Qy	1863	ACAAACAGCACCATGATGTATGTTCGGTGGTTTCAATATGTCCTCTCCTCAGCGACATCTCT	1922
Db	1929	GTAAACAGCACCATGATGTATGTTCGGCGGCTTCAACAGCCTCTCTCTCAGTGACGTCTTT	1988
Qy	1923	GGTATTCACCTCGAACAGTGTGATGGCATCGGAGTGAAGCGCTGTTTATACACAGG	1982
Db	1989	GGTCTTTACCTCGGAGCAGTGGCATGCAACCCGACGTGAAGCTGCTTTGTGTGCAGCAGG	2048
Qy	1983	ACCTGGTATTCGGTGTGTGTGGAACACAGGGTCTGTCTCAGTGTATCTCGTGGCGCTGGC	2042
Db	2049	ACCTGGTATTCGGTGTGTGTGGGACACACAGTCTGTCTCGATGTACCTCTCTGGAGTTGGC	2108
Qy	2043	AACGTATGAACAGAGAAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCA	2102
Db	2109	AACGTGAAGAACAAAGCAGAAAAGTTAAATCAGAGTGTTTTTCTTAAAGAAACCCCTTGACCA	2168
Qy	2103	TGACAGATGTGACACAGACAGATTGTTACAGCTGTACAGCCAAACCAATGACATGGCCA	2162
Db	2169	TGACAGATGTGACAGACACAGATTGTTACAGCTGCACAGCCAAATACCAATGACTGGCCA	2228
Qy	2163	CTGGTGCAATGCAACATTGTGTCCCGAGAACACACAGCTGTCTAGAGGCCCAGATCTCCAT	2222
Db	2229	CTGGTGCAATGCAATCACTGTCTCTGTGAACACACAGCTGCACAGAGGCCCAGATCTCCAT	2288
Qy	2223	TTTTAGGTATGAGNAATGCCCCCAGAGNATACCCCATGTACTCTGTATAACAGAGACCCAG	2282
Db	2289	TGCCAAGTATGAGAGTTGCCCCCAGAGNATACCCCATGTACTCTGTCAATTAAGAAACCCAG	2348
Qy	2283	CTGCAGAGCTGTGCCCTGGACACAGAACTGCCAGTGGAGCCCCCGGAATCAGAGTGCAT	2342
Db	2349	CTGCAGAGCTGTGCCCTAGACCAGAACTGCCAGTGGAGCCCCCGGAATCAAGAGTGCAT	2408
Qy	2343	TGCCCTGCCCGAAAATATCTGTGGCAATGGCTGGCAATTTGGTTGGAACCTCATGTTTTGAA	2402
Db	2409	CGCCCTGCCCGAAAATATCTGTGGCAATGGCTGGCAATTTGGTTGGAACCTCGTGTCTGAA	2468
Qy	2403	AATTACTACTGCCAAGAGGAATATAGCAATGCTTAAATGCTTCTGTAGGAACACCAATGTC	2462
Db	2469	AATCACTACTGCTAAGAGGAATATAGCAATGCTTAAATGCTTCTGTAGGAACACCAATGTC	2528
Qy	2463	CTTTTGGCTTCTCTTAAACCCAGAGAAGGTAGAAATTTGTCTCTTAAGCAGCTGCGAAT	2522
Db	2529	CTTTTGGCTTCTCTCTCATCCCAAGAGAAGGTGGAGTTTGTCTCTTAAGCAGCTTCGATT	2588
Qy	2523	AATGAGTCACTCTCAGAGCATGTCCAGCTCACTTAAACCCCATGGGTGCGCCTTCGAA	2582
Db	2589	AATGCAATCATCTCAAGATATGTCCAGAGCTCACTCTGACTCCATGGGTGTGCTCTTCGAA	2648
Qy	2583	GATCAATGTCTCTACTGGTGTCTGGGAAGATATCTCCCACTTTTCAAAATAGTTTACTACA	2642
Db	2649	GATCAATGTCTCTACTGGTGTCTGGGAAGATATGTCCTCATTCACAAATAGTTTGTCTGCA	2708
Qy	2643	GTGAGTCCGCTCTCAGCCCACTGATGCTGGATCTCTGTGGANTTTTATCAGAACCCAGTAC	2702
Db	2709	GTGAGTCCATCTCAGCCCACTGATGCTGGCTCTCTGTGGATCTTGTGAGAGCTAGTAC	2768











```

/strain="WTC"
/db_xref="taxon:10116"
1. .8739
/gene="atrn"
61. .4359
/gene="atrn"
/codon_start=1
/product="attractin"
/protein_id="BAB21017.1"
/db_xref="GI:1275308"
/translation="MVAANAATAEARLGRYTTATAAPGAWKERQHRPCAATGAWRP
WPRAGLCLPVLSSALSPPLPLLPLLLPLPREAANAATAAASVGSAAASGAEK
ECDPCVNGCRNCPGTGOCVCTGCGQCHGGRFLRTGSGSVFTDGGNVAYVYK
CTWLLIEQPKIMELRLRNPHATECSDWHLVYDGDLSIYAPLIAAFSLGLVPERDGNET
APEVITSGVALLHGFSDAANLTFNITNFDMCPNNSCGRGECSSNSSSVYCEGC
SENKWSGCDIPHCTDNCGRPHRIGVICNASDTRGCGFPFHWQGGPCCSIYPVANSQWTR
EYGDGLKPRASHKAVNGNIMVYGGVFNHSDYSVMVLAIDLASREWLSNHSVNSV
VVRGCHSLALHKDKIDSTGNVTELRFVFIHNESWLLTLPKADQVQGGYHSS
SAHVTLSGRGVMVLVYFGHCLPYGYISVYQVELEKNTWSLIQTQALVQGGYGHSS
VYDHTKALYVHGCVYSGHAKFADLADLYRVHVDTCMTILKDSRPFRLVHTAVIUS
GTMVLVFGTNGTDLTSMGSHGKCFSSDEFMAYDIAQCRWSVLPRPELHDHVRFGHSALV
HNSYVYVPGFNLSILLSDVLVFTSEQCDARHSEAAVYAGPGIRCLVDCQTSSTSGE
LATBEQAKDKSECFKRTLDHDKDQHTDCSTANTINDCHWCNDRVFNHSCSTSGE
QISIAQYDNCDFNPMYCNKTKSCRSCLADQNCQWEPBQEOCALPENICGLHWHLV
GNSCLTKITAKENTDNKLSCRNENAFILSTSQKVEFVYLKQLRMSQSSKULTIN
TPWGLKRIYVNCWEDMSFTNLSLQWMPSEPSDFAGFCGILSEPTSGRLKAATCIN
PLNGSVETPANHSAKCCPCALTACGECTSSSCECMNCKQSDNNAVYVAGFP
FQGCWERYPSSAPKENCSCYCTCSHGLEOPGCGWCTDPSNTGKGKCIEGSYGPVKM
PASHATGNVYQPLLNSMCLSDSRYNWSIHCPCACQNGKSCIKNOSICEKEDLIT
GHKCECTISGFGYDTGKQCKPCMGHASLNCNTNKGCTCTYKGVKGEBCOLCEVE
NRYGNPLKGYCYTLLIDYQFTLSQEDDXYTIANVATPDQNRDLMDFINASK
NFENLNIWATSPFAGTQDGEVPVYVSKTNIKEYKDSFSNEKFPFRNHPNITFFVYVNN
FTWPIKIQISQHSNFDLVQFFVTFSCFLSILLVAAYVMKIKQSCWASRRRQELL
REMOQMASRPSAVNALDEEPPDLIGSITKVPKPIALEPCFGKNAVLVSFVL
BEGGLTPEPCGSIAYASALVNI SQOMPIVWFKSGAVNRKQOOPAQPGTIC"

```

## ORIGIN

Query Match	81.6%;	Score 3114.8;	DB 10;	Length 8739;
Best Local Similarity	88.8%;	Pred. No. 0;		
Matches 3371; Conservative	0;	Mismatches 427;	Indels 0;	Gaps 0;
QY	3	GGTGGCCGACGGCGGCAACTGAGGCAAGGCTGAGGAGGAGACGGCGCGACGGAGC	62	
Db	72	GGCGGCGCGGAGGCGACTGAGGCGCGGTTGAGGGGTACACGACGGCGACAGCAGC	131	
QY	63	GCTGCGGGGACAGACGGCGGCGGCGACTGGGACTGGGACGTGACACAGGGCTGGAGGCC	122	
Db	132	GCCCGCGGGCTGGAAGGAGGACGACCCGACCCCTGCGCGCGGACAGGGGGCTTGGAGGCC	191	
QY	123	GGGGCTGGGGGCGGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGTGTGGGCCACAGGCT	182	
Db	192	GTGGCGCGCGCGCGGGCTGTCTCCCGGGTGTGTTCGCGGGCTGTATCCCGCGCGCC	251	
QY	183	GCTGCTGCTGTTGTTTGTCTCCCGCGCGCGCTGTTCGTGCTGTCTGTCCTGTGGAGC	242	
Db	252	GCTGCTCGCGTGTGCGGCTGTCTTTTCGCTATTGCTGTGCACATGCCCCGGGAGC	311	
QY	243	CGAGGCGCGCGCGCGCGCGGGTGTGGGGCTCAGCCCGACGCGAGGCGCAAGGAATG	302	
Db	312	CGAGGCGCGTGCCTGGCGGCGGGTGCTCAGGCTCGSCCGCAGCTGAGGCCAAGGAATG	371	
QY	303	TGACCGGCGCTGTGTCAACCGCGGTCGTTGCAACCTTGGACCGGCGCAGTGCCTTCGCC	362	
Db	372	TGACCGGCGCGTGTGTCAACCGCGCGCGGTGCAACCCCTGGCACCGGCGAGTGCCTTCGCC	431	
QY	363	CGCCGCTGGGTGGGCGGACGAATGCACGACTGCGGGGGCGCGCTTCAGACTAACTGGATC	422	
Db	432	CACGGGCTTGGTGGGCGGAGCAATGCCAGCACTGCGGGGGCGGCTTCAGACTAACTGGATC	491	
QY	423	TTCTGGGTTTGTGACAGATGGAACCTGGAAATTTATAATACAAAACGAAGTGCACGTGGCT	482	
Db	492	TTCTGGGATTTGTAAACAGATGCACTGGGAATTTATAATATAAAACGAATGCACATGGCT	551	

QY	483	CA	TTGAAGGACGACCAATATAGAAATTAATGAGACTTCGTTTCAATCAATTTTGCTACAGATG	542
Db	552	CA	TTGAAGGACGACCAATTAAGATTAATGAGACTTCGATTCACCAATTTTGCTACTGAATG	611
QY	543	TAG	TTGGGACCAATTTATATGTTTATGATGGGACACTCAATTTATGACACCGCTAGTTGCTGC	602
Db	612	TAG	CTGGGACCAATTTATATGTTTATGATGGGACACTCAATTTACGACACCTCTGATTTGCTGC	671
QY	603	A	TTTAGTGGCCCTCATGTTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTCG	662
Db	672	C	TTTAGTGGCCCTCATGTTTCCTGGAAGAGATGGCAATGAGACGGGCCCTGAACTCACTGT	731
QY	663	CA	CATCAGGTTATGCTTCCTGCTGCATTTTTTTAGTCATGCTGCTTATATAATTTCACTGGATT	722
Db	732	CA	CTTCAGGTTATGCACTACTGCAATTTTTTTCACTGATGCTGTTTATATCTGACTGGATT	791
QY	723	T	AATATTAATCTTACAGTTTGTATATGTGTCACAAATAACTGCTCAGGCCGAGGAGAGTGTAA	782
Db	792	T	AATATCACTTACAAATTTTGACATGTGCCGGAATAATTTGCTCAGGCCGAGGAGAGATGTA	851
QY	783	G	ATCAGTATACAGCGAAACTGTTTCAATGTCAATGTTCTGAAAACTCGAAAAGGTGAAGC	842
Db	852	G	AGCAGTAACAGCAGCAGCACCTGTTTGAGTGTGAATGTTCTGAAAACTCGAAAAGGGAGTC	911
QY	843	A	TGTGACATTCCTCACTGTACAGACAACTGTGGTTTTTCTCTCATCGAGGCATCTGCAATTC	902
Db	912	C	TGTGACATTCCTCACTGTACAGACAACTGTGGTTTTTCTCTCACCGAGGCATCTGTAAACGC	971
QY	903	A	AGTGTATGTACAGGATGCTCTGCTTCTCAGACTGGCAGGGTCTGGAATGTTTCAGTCC	962
Db	972	A	AGTGCACACAGAGGCTCTCTGCTTCCTCACTGGCAGGGTCTGGAATGTTTCAATTC	1031
QY	963	T	GTACCAGCTAACCACTCATTTTGGACTCGAGAGGAATTTCTAACTTAAAGCTCCCCAG	1022
Db	1032	T	GTGCCAGCTAACCACTCTTTTGGACTCGAGAGGAATTTCTGACTTAAAGCTCCCCAG	1091
QY	1023	A	GCATCTCTATAAGCTGTGGTCAATGGAACATTAATGTGGGTTGTTTGGAGGATATATGTT	1082
Db	1092	A	GGCTCTCTATAAGCTGTGGTCAATGGAACATTAATGTGGGTTGTTTGGCGGGTATATGTT	1151
QY	1083	C	ACCACTCAGATTATAACATGGTTCTACGATATGACCTTGCTTCTAGGAGATGGCTTCC	1142
Db	1152	C	ACCACTTCAATTAACAGATGGTTTTTACGATATGACCTGGCTTCTAGGGAATGGCTTTC	1211
QY	1143	A	CTAAACCGTCTCTGAAACAAATGTGGTGTGATAGATGTCATCTTTTGGCATTAACAA	1202
Db	1212	A	CTAAACCACTCTCTGAAACAGTGTGGTGTGAAATATGGTCAATCTTTGGCATTAACAA	1271
QY	1203	G	ATAAAATTTACATGTATGGAGGAATAATGATCCAACTGGGAATGTGACCAATCAGTT	1262
Db	1272	G	ATAAAATTTACATGTATGGAGGAATAATGATCCAACTGGGAATGTGACCAACAGCT	1331
QY	1263	G	AGAGTTTTTTCATTAATCAGTCACTGGGTGTTTGTGACCCCTTAAGCCAAAGAGCA	1322
Db	1332	G	AGAGTATTTTCATTAATCAATATCAATCATGGGTATTGTTAAACCCCAAGCTAAGGATCA	1391
QY	1323	G	TATGAGTGGTTTGGGCACTCTGCACACATTTGTTTACCTGAGAAATGGCCGAGTGTCA	1382
Db	1392	G	TATGAGTGGTTTGGGCACTCAGCACACATTTGTTTACCTGTCACTGGCCGCTGTGGTCA	1451
QY	1383	G	CTGGTTCATCTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGA	1442
Db	1452	G	TTGGTTCATCTTTGGTCACTTTGTCGGCTCTACGGATATATAAGTGTGTGCAGGAATATGA	1511
QY	1443	T	TGTGGATAAGAACACATGGAGTATATTAACACCCAGGGTGCCTTGTGCAAGGGGGTTA	1502
Db	1512	C	TTTGGAAAACACATGGAGTATATTAACAACTCAAGGTGCTCTAGTGCAGAGGGGGTTA	1571
QY	1503	C	GGCCATAGCAGTGTTTACGACCATAGGACCGGCCCTTATACGTTTCATGGTGGCTACAA	1562
Db	1572	T	GGCCACAGCAGTGTTTATGATCAGAGCAACAGGCTCTGTATGTTCATGGTGGCTACAA	1631
QY	1563	G	GGCTTTTCAAGTGAACCAAGTACCGGGTTGCAGATGATCTCTACCGATATATGTGGATAC	1622

Db 1632 GGCCTTTTCAGCGCCAAACAGTACCGGCTTGCGATGACCTCTACAGATACCAATGTGGACAC 1691  
QY 1623 CCAGATGTGGACCATCTTTAAAGACAGCGGATTTTCGGTTACTTTCCGACACAGCTGTGAT 1682  
Db 1692 TCAGATGTGGACCATCTTTAAAGACAGCGGATTTTCGGTTACTTTCCGATACAGCTGTGAT 1751  
QY 1683 AGTGAAGTGAACCATCTGCTGGTGTGGGGGAAACACACACAAATGACATCTATCAGGCA 1742  
Db 1752 AGTGAAGTGAACCATCTGCTGGTGTGGGGGAAACACACACAAATGACATCTTCCATGAGCA 1811  
QY 1743 TGGCGCCAAATGCTTCTCTCAGATTTCAATGCTTATGAGCTTATGACATGCTGACCGCTGTGC 1802  
Db 1812 CGGTGCGAAATGCTTCTCTCAGATTTCAATGCTTATGACATTTGCTGACCGGATGTC 1871  
QY 1803 AGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGAGCTCT 1862  
Db 1872 AGTCTCTCCAGACCTGAGCTCCATCAGATGTCAACCGATTTGGCCACTCAGAGCTCT 1931  
QY 1863 ACACAAAGACACATGATGTGTTTCGGTGGTTTCAATAGTCTCTCTCAGCGACATCCT 1922  
Db 1932 GCACAAAGACACATGATGTGTTTCGGTGGTTTCAACAGCTCTCTCTCAGTGAAGCTCT 1991  
QY 1923 GGTATTTCACTCGGAACAGTGTGATGCGCATCGGAGTGAAGCGCTTCTTTAGCAGCAGG 1982  
Db 1992 AGTCTTCACTCGGAGCAGTGTGATGCAACCGCAGCGAGAGCTGCTTGTGTGGCAGCAGG 2051  
QY 1983 ACTCGATTTCCGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGC 2042  
Db 2052 ACTCGCATCCGCTGTGTGTGGGACACACAGTCGTCTCGATGTACTCTCTGGAGTTGGC 2111  
QY 2043 RACTGATGAACAGAGAAAGTTAAATCAGATGTTTTCBAAAGAACTCTTGACCA 2102  
Db 2112 AACTGAAGAACAGCAGAAAGTTAAATCAGATGTTTCTTAAAGAACCTTGACCA 2171  
QY 2103 TGACAGATGTGACACAGCACACAGATTTGTACAGCTGTACAGCCAAACACAAATGACTGCCA 2162  
Db 2172 TGACAGATGTGACACAGCACACAGATTTGTACAGCTGTACAGCCAAATACCAATGACTGCCA 2231  
QY 2163 CTGGTGCAATGACATTTGTGTCGCCAGGAAACACAGCTGTCTCAGAGCCAGATCTCGAT 2222  
Db 2232 CTGGTGCAATGATCACTGTGTCTTCTGTAACACACAGCTGTGACAGAGCCAGATCTCCAT 2291  
QY 2223 TTTTGTATGAGATTTGCCCAAGCATAAACCCATGTAAGTGTAACTGTAACAGAGACAG 2282  
Db 2292 TGCCAGTATGACATTTGCCCAAGCATAAACCCATGTAAGTGTAACTGTAACAGAGACAG 2351  
QY 2283 CTGAGGAGCTGTGCTTGGACAGAACTGCGAGTGGAGCCCGGAAATCAGGAGTGCAT 2342  
Db 2352 CTGAGGAGCTGTGCTTGGACAGAACTGCGAGTGGAGCCCGGAAATCAGGAGTGCAT 2411  
QY 2343 TGCCCTGCGCGAAATATCTGTGCGATTTGGCTGCGATTTGGTGGAACTCATGTTTGA 2402  
Db 2412 CGCCCTGCGCGAAATATCTGTGCGATTTGGCTGCGATTTGGTGGAACTCATGTTTGA 2471  
QY 2403 AATTACTACTGCGAGGAGATTTATGCAATGCTAAATTTGCTGTAGGAAACCAATGC 2462  
Db 2472 AATCACTACTGTAAGGAAATTTATGCAATGCTAAATTTGCTGTAGGAAACCAATGC 2531  
QY 2463 CTTTTTGGCTTCTCTTACAAACCCAGAGAGGTAGATTTGCTCTTAAAGCGCTGCGAAT 2522  
Db 2532 CTTTTTGGCTTCTCTTACAAACCCAGAGAGGTAGATTTGCTCTTAAAGCGCTTGCAT 2591  
QY 2523 AATGAGTCTATCTCAGAGCATGTCCAGCTCACTCTGACTGCTGCTTGGCTTGGAA 2582  
Db 2592 AATGCAATCACTCAGAGCATGTCCAGCTCACTCTGACTGCTGCTTGGCTTGGAA 2651  
QY 2583 GATCAATGTCTCTACTGCTGCGAGAGATATGCTCCCATTTACAAATAGTTTACTACA 2642  
Db 2652 GATCAATGTCTCTACTGCTGCGAGAGATATGCTCCCATTTACAAATAGTTTGGTGA 2711  
QY 2643 GTGGATGCGCTCTGAGCCAGTGTGCTGGATTTCTGTGGATTTTATCAGAACCCAGTAC 2702

Db 2712 GTGGATGCCATCTGAGCCAGTGTGCGGCTTCTGTGGATCTTGTGAGCGCTAGTAC 2771  
QY 2703 TCGGGGACTCAAGGCTGCAACCTGCACTCAACCACTCAATGGTAGTGTGTGAAAGGCC 2762  
Db 2772 TCGGGGACTCAAGGCTGCAACCTGCACTCAACCACTCAATGGTAGTGTGTGAAAGGCC 2831  
QY 2763 TGCAACCAACAGTGTCTAAGCAGTCCCGGACACCATGTGCTTTGAGGACAGCATGTGGAGA 2822  
Db 2832 TGCAACCAACAGTGTCTAAGCAGTCCCGGACACCATGTGCTTTGAGGACAGCATGTGGAGA 2891  
QY 2823 TTGCACACAGCGGAGCTCTGAGTGTGCTGTCAGACAAATGAGACAGTGTGTGACTC 2882  
Db 2892 GTGCACACAGCAGCTCCGAATGCAATGTGGTGCAGCAACATGAGCAGTGTGGTAGACTC 2951  
QY 2883 CAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGATATACCATGAGCAC 2942  
Db 2952 CAATGCTATGTAGCTCTCTTCTTTCGGCCAGTGTATGGAGTGTATACATGAGCAG 3011  
QY 2943 CTGCCCCCTGAAAAATTTGTTACAGCTTACTGTACTGTAGTCAATTTGCTTGGAGCAACGAG 3002  
Db 3012 CTGCCCCCTGAAAAATTTGTTACAGCTTACTGTACTGTAGTCAATTTGCTTGGAGCAGCAGG 3071  
QY 3003 CTGTGGCTGTGTACTGATCCAGCAATACTGSCAAAGGAAATGATAGAGGTTTCTTA 3062  
Db 3072 CTGTGGTGTGTACTGATCTTACATCTAGCAATTTGGAAGGCAATGTATTTGAGGAGCTA 3131  
QY 3063 TAAAGACAGTGAAGATGCTTTCGACGCCCTACAGGCCCTATCCAGAAATTTCTATCCAGCCCT 3122  
Db 3132 TAAAGACAGCTGGAAGATGCTTTCATGCTCTACAGGAAATTTGTTACCCACAGCCCT 3191  
QY 3123 GCTCAATTTCCAGCATGTGTAGAGACAGCAGATACATCTGCTTTTCACTTCACTGTCC 3182  
Db 3192 TCTGAATCTCAGCATGTGTGTAGAGACAGCAGATACATCTGCTTTTATTTCACTGTCC 3251  
QY 3183 AGTTTCCCAATGCAACCGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGA 3242  
Db 3252 AGTTTCCCAATGCAATGGAACAGCAAGTGTCAATCAACAGAGCATCTGTGAGAGTGTGA 3311  
QY 3243 GAACCTGACACAGGCAAGCTGTGAGACCTGTGCAATCTGCTTCTACGCTGTACGCTGATCCAC 3302  
Db 3312 GGACCTGACACAGGCAAGCTGTGAGACCTGTGCAATCTGCTTCTATGCTGAGCCGAC 3371  
QY 3303 CAATGAGGAAATGTCAAGCTGCAATGCAATGGGCAAGCTCTGTGTGCAACACCAA 3362  
Db 3372 TAATGAGGCAATGTCAAGCTGCAATGGGCAAGCTCTGTGTGCAACACCAA 3431  
QY 3363 CAGGCAAGTGTCTGTGACCAACCAAGGCTCAAGGGGAGAGTGTGCTAGCTATGTGA 3422  
Db 3432 CACTGCAAGTGTCTGTGTACCAACCAAGGCTCAAGGGGAGAGTGTGCTAGCTATGTGA 3491  
QY 3423 GGTAGAAATTCATACCAAGGAAACCTCTCAGAGAAACATGTTATTTACTTCTTAT 3482  
Db 3492 GGTAGAAATTCATACCAAGGAAACCTCTCAGAGAAACATGTTATTTACTTCTTAT 3551  
QY 3483 TGACTATCAGTTTCACTTTTAGCTTATCCAGAGATGATCGCTATTTACAGCTATCAA 3542  
Db 3552 TGACTATCAGTTTCACTTTTAGCTTATCCAGAGATGATCGCTATTTACAGCTATCAA 3611  
QY 3543 TTTTGTGGCTTCTCTGAGCAACAAACAGGATTTGGACATGTTTCAATGCTTCTTAA 3602  
Db 3612 TTTTGTGGCTTCTCTGAGCAACAAACAGGATTTGGACATGTTTCAATGCTTCTTAA 3671  
QY 3603 GAATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3662  
Db 3672 AAATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGCTGGAGA 3731  
QY 3663 AGAGATGCTGTGTTTCAAAACCAACATTTAGAGAGTACAAAGATGTTTCTCTAATGA 3722  
Db 3732 AGAGTGTGCTGTGTTTCAAAACCAACATTTAGAGAGTACAAAGATGTTTCTCTAATGA 3791  
QY 3723 GAAGTTTGTATTTTCCCAACCCCAATATCACTTTCTTTTGTATGTCAGTAATTTTCAAC 3782  
Db 3792 GAATTTGTATTTTCCCAACCCCAATATCACTTTCTTTTGTATGTCAGTAATTTTCAAC 3851



Db 1092 AGCCTCTCAATAAGCTGAGTCAATGGAAATATAATGTGGTGTGGCGGTATATGTT 1151  
QY 1083 CAACCACTCAGATTATAAATGTTCTAGCTATGACCTTGTCTTAGGGAGTGGCTTCC 1142  
Db 1152 CAACCACTCAGATTACAGATGTTTACGCTATGACCTGGCTTCTAGGGAATGGCTTTC 1211  
QY 1143 ACTAAACCGTCTGTGAACAAATGTGGTGTGTAGATATGGTCAATCTTTGGCAATATACAA 1202  
Db 1212 ACTAAACCAATCTGTGAACAGGTGTGTGTAAAGATATGTCATCTTTGGCAATATACAA 1271  
QY 1203 GGATAAAATTTACATGTATGGAGGAATTTGATCCNACTGGGAATGTGACCAATGAGTT 1262  
Db 1272 GGATAAAATTTACATGTATGGAGGAATTTGATCCNACTGGGAATGTGACCAAGAGCT 1331  
QY 1263 GAGAGTTTTCATATTCATAATGATCAATGGTGTGTGACCCCTAAGGCAAGAGGACA 1322  
Db 1332 GAGAGTATTTTCATATTCATAATGATCAATGGTGTGTGACCCCTAAGGCAATGAGTCA 1391  
QY 1323 GTATGAGTGGTGGCACTCTGACACATTTGTTACATGTAAGAAATGGCCGAGTGGTCAAT 1382  
Db 1392 GTATGAGTGGTGGCACTCTGACACATTTGTTACATGTAAGAAATGGCCGAGTGGTCAAT 1451  
QY 1383 GCTGGTCACTTTGGTCACTGCGCTCTCTATGGATATATAGCAATGTGACAGGAATATGA 1442  
Db 1452 GTTGGTCACTTTGGTCACTGCGCTCTCTATGGATATATAGCAATGTGACAGGAATATGA 1511  
QY 1443 TTTGGATAGAACATGAGATATATACACACCCAGGCTGCTTGTGCAAGGGGTTA 1502  
Db 1512 CTTGGAAAGAACACATGGAGTATATACAACTCAAGGTGTCTTAGTGCAAGGGGTTA 1571  
QY 1503 CGGCCATACAGTGTGTACGACCATAGGACAGGCGCTTATAGTTCATGTGTGCTACAA 1562  
Db 1572 TGGCCACAGCATGTTTATGATCAGGACCAAGGCTCTGATGTTCAATGGTGGCTACAA 1631  
QY 1563 GCGTTTCAAGTGCATAATAGTACCGCTTGCAGATGATCTTACCGATATGATGTGGATAC 1622  
Db 1632 GCGTTTCAAGGCAACAAAGTACCGCTTGCAGATGATCTTACAGATACCAATGTGGACAC 1691  
QY 1623 CCAGATGTGGACCATCTTAGGACAGCGGATTTTCGTTACTTGTGACACAGCTGTGAT 1682  
Db 1692 TCAGATGTGGACCATCTTAGGACAGCGGATTTTCGTTACTTGTGACATACAGCTGTGAT 1751  
QY 1683 AGTGAAGTGAACCATGCTGTGTGTGGGGGAAACACACAAATGACACATCTATGAGCCA 1742  
Db 1752 AGTGAAGTGAACCATGCTGTGTGTGGGGGAAACACACAAATGACACATCTATGAGCCA 1811  
QY 1743 TGGCCCAAAATGCTTCTCTCAGATTTATGCGCTATGACATTTGCTGTGACCGCTGGTC 1802  
Db 1812 CGGTGCCAAATGCTTCTCTCAGATTTATGCGCTATGACATTTGCTGTGACCGCATGGTC 1871  
QY 1803 AGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGCGCCATTCAGCAGTCTT 1862  
Db 1872 AGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGCGCCATTCAGCAGTCTT 1931  
QY 1863 ACACAAAGCAACCATGATGTGTGGTGTTCATAGTGTCTCTCTCAGCGCATCTCT 1922  
Db 1932 GCACAAAGCAACCATGATGTGTGGTGTTCATAGTGTCTCTCTCAGCGCATCTCT 1991  
QY 1923 GGTATTCACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCGCTGTTTATGACGAGG 1982  
Db 1992 AGTCTTCACTCGGAGCAGTGTGATGACACCGCAGGAAGCTGCTTGTGTGGCAGCAGG 2051  
QY 1983 ACCTGGTATTCGGTGTGTGGAAACACAGGCTCGTCTCAGTGTATCTCGTGGCGCTGGC 2042  
Db 2052 ACCTGGCATCCGGTGTCTGTGGGACACACAGTGTCTCGATGTACCTCTCGGAGTGGC 2111  
QY 2043 AACTGTAGAACAGAGAAAGTTAAATCAGATGTGTTTTCAAAAGAACTCTTGACCA 2102  
Db 2112 AACTGAAGAAACAGGAGAAAGTTAAATCAGATGTGTTTCTTAAAGAAACCTTTGACCA 2171  
QY 2103 TGACAGATGTGACGAGCACAGATTTGTTACAGCTGTACAGCCAAACCAATGACTGCCA 2162

Db 2172 TGACAGATGTGACGAGCACACAGATTTGTACAGCTGCACAGCCAAATACCAATGACTGCCA 2231  
QY 2163 CTGGTCAATGACCATTTGTCTCCAGGAAACACAGCTGCTCAGAGGCCAGATCTCCAT 2222  
Db 2232 CTGGTCAATGATGATCAGTGTCTCCAGGAAACACAGCTGCTCAGAGGCCAGATCTCCAT 2291  
QY 2223 TTTTGTATGAGAAATTTGCCCAAGGATAACCCCATGTACTCTGTAAACAAGAACCCAG 2282  
Db 2292 TGCCAGATATGACAAATTTGCCCAAGGATAACCCCATGTACTCTGTAAACAAGAACCCAG 2351  
QY 2283 CTGCAGAGCTGTGCTCCAGGAAACACAGCTGCTCAGAGGCCAGATCTCCAT 2342  
Db 2352 CTGCAGAGCTGTGCTCCAGGAAACACAGCTGCTCAGAGGCCAGATCTCCAT 2411  
QY 2343 TGGCCTGCCCAAAATATCTGTGCAATTTGGTGGCAATTTGGTGGAAATCTCATGTTTGA 2402  
Db 2412 CGCCTGCCCAAAACATCTGTGCAATTTGGTGGCAATTTGGTGGAAATCTCATGTTTGA 2471  
QY 2403 AATTACTATGCGCAAGGAGAAATTTATGACAAATGCTAAATTTGTTCTGAGGAAACCAATGC 2462  
Db 2472 AATCACTACTGCTAAGGAAATTTATGACAAATGCTAAATTTGTTCTGAGGAAACCAATGC 2531  
QY 2463 CTTTGTGGCTTCTTTACAAACCCAGAGAGGTAGAAATTTGCTTAAAGCAGCTCGAAT 2522  
Db 2532 CTTTGTGGCTTCTTTACAAACCCAGAGAGGTAGAAATTTGCTTAAAGCAGCTCGAAT 2591  
QY 2523 AATGCACTCATCTCAGAGCAATGCTCAAGCTCACTTAAACCCAGGCTCGGCTTGGAA 2582  
Db 2592 AATGCACTCATCTCAGAGCAATGCTCAAGCTCACTTAAACCCAGGCTCGGCTTGGAA 2651  
QY 2583 GATCAATGTCTCTAGTGTGGGAGATATGCTCCCAATTTACAAATGTTTACTACA 2642  
Db 2652 GATCAATGTCTCTAGTGTGGGAGATATGCTCCCAATTTACAAATGTTTACTACA 2711  
QY 2643 GTGGATGCGCTGTGAGCCCAAGTATGCTGGATTTCTGTGAAATTTTATCAGAAACCCAGTAC 2702  
Db 2712 GTGGATGCGCTGTGAGCCCAAGTATGCTGGATTTCTGTGGATTTCTCAGAGCCCTAGTAC 2771  
QY 2703 TCGGGAGCTGAGGCTGCAACCTGATCAACCCATCAATGGTGTGCTGTGAAAGGCC 2762  
Db 2772 TCGGGAGCTGAGGCTGCAACCTGATCAACCCATCAATGGTGTGCTGTGAAAGGCC 2831  
QY 2763 TGCAAAACCAAGTGTAAAGCAGTGCAGGACACATGCTGCTTGGAGCAGAGCTGTGAGA 2822  
Db 2832 TGCAAAACCAAGTGTAAAGCAGTGCAGGACACATGCTGCTTGGAGCAGAGCTGTGAGA 2891  
QY 2823 TTGCAACCAAGGAGCTGTGAGTGTGATGTGTCAGCAACATGAAGCAGTGTGAGCTC 2882  
Db 2892 GTGCAACCAAGGAGCTGTGAGTGTGATGTGTCAGCAACATGAAGCAGTGTGAGCTC 2951  
QY 2883 CAATGCTGTGTGCTGCTTCCCTTGGCCAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTA 2942  
Db 2952 CAATGCTGTGTGCTGCTTCCCTTGGCCAGTGTGTAAGTGTGTAAGTGTGTAAGTGTGTA 3011  
QY 2943 CTGCCCCCTGAAAATTTGTTGAGGCTACTGTACTGTACTGTACTGTCTGAGCAACCCAGG 3002  
Db 3012 CTGCCCCCTGAAAATTTGTTGAGGCTACTGTACTGTACTGTACTGTCTGAGCAACCCAGG 3071  
QY 3003 CTGTGCTGTGTGATGATGCTCCAGCAATCTGTCAGGAAAGGAAATGATAGAGGTTCTTA 3062  
Db 3072 CTGTGCTGTGTGATGATGCTCCAGCAATCTGTCAGGAAAGGAAATGATAGAGGTTCTTA 3131  
QY 3063 TAAAGGACCAAGTGAAGTGTCTTCCAAAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3122  
Db 3132 TAAAGGACCAAGTGAAGTGTCTTCCAAAGCCCTTACAGGAAATTTCTATCCACAGCCCT 3191  
QY 3123 GCTCAATTCAGCATGTGTGTAGAGGACAGGAGATACAACTGGTCTTTTCAATCTACTGTCC 3182  
Db 3192 TCTGAATCCAGCATGTGTGTAGAGGACAGGAGATACAACTGGTCTTTTATTCATCTGTCC 3251  
QY 3183 AGCTTGCCAATCAACCGGCAAGT 3242  
Db 3252 AGCTTGCAGT 3311



QY	483	CATTGAAGCAGCAGCCAAATAGAAATAGAGACTTCGGTTTCAATCATTTTGTCTACAGATG	543
Db	506	CATTGAAGGCATATCCAAATGCAGTCTAAGGTTAAGATTCAACCAATTTTGCTACAGAATG	565
QY	543	TAGTTGGGACCAATTTATATATGTTTATGATGGGACCTCAATTTATGCAACGCTAGTTGCTGC	602
Db	566	TAGCTGGGACCAATTTATATATGTTTATGATGGGACCTCAATCTACGCACCTCTGATTGCTGC	625
QY	603	ATTTAGTGGCCTCATTTGTTCTTGACAGAGATGGCAATGAGACTGTCCTTGAGGTGTTTGC	662
Db	626	CTTTAGTGGCCTCATTTGTTCTGAAAGAGATGGCAATGAGACGGCTCTCTGAGGTCACTGT	685
QY	663	CACATCAGGTTTANGCCTTGCTGTCATTTTTTTTAGTGATGCTGCTTATAATTTGACATG	722
Db	686	CACITTCAGGTTATGCACTGCTGTCATTTTTTTCAGTGATGCTGCTTATAATCTGACATG	745
QY	723	TAAATTACTTACAGTITTTGATATGTCGCCAAATAACTGCTCAGCCGACGAGAGATGTAA	782
Db	746	TAAATCACTTACATNTTTGACATGTGTCCGAAATTAATTTGCTCAGCCCGAGAGAGATGTAA	805
QY	783	GATCAGTAATAGCAGCGAAACTGTTGAAATGTGAATGTTTCTCAAACTCGAAAGGTGAAGC	842
Db	806	GAGCAGTAAACAGCAGCAGCGCTGTTGAGTGTGAATGTTTCTCAAACTCGAAAGGGAGTC	865
QY	843	ATGTGACATTCCTCACTCTACAGACAACATGTGTGTTTTCCTCATCGAGGCATCTCGAATTC	902
Db	866	GTGTGACATTCCTCACTCTACAGACAACATGTGTGTTTTCCTCACCAGAGCATCTGTAATGC	925
QY	903	AAGTGATGTCAGAGGATCTCCTGCTTCTCAGACTGGCAGGGTCTCTGATGTTTCAGTTCC	962
Db	926	AAGGATACCAAGAGGGTCTCCTGCTTCTCACTGGCAGGGTCTCTGATCTCAATTTCC	985
QY	963	TGTACCAGCTAACCAAGTCATTTTTGGACTCGAGAGGAATTTCTAACTTAAAGCTCCCCAG	1022
Db	986	TGTGCCAGCTAACCAAGTCATTTTTGGACTCGAGAGGAATTTCTCATTTAAAGCTTCCCG	1045
QY	1023	AGCATCTCATAAAGCTGTGGTCAATGGAAACAATATGTGGGTTCTGTGGAGGATATATGTT	1082
Db	1046	AGCCTCTCATAAAGCTGTGGTCAATGGAAATAATATGTGGGTTCTGTGGCGATATATGTT	1105
QY	1083	CAACCACTCAGATTTAATACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTC	1142
Db	1106	CAACCATTCAGATTCAGCATGGTTTGTAGCGTATGACCTGACTGACTTCTAGGGAAATGGCTTC	1165
QY	1143	ACTAAACCGTTCTGTGAACAATGTGGTTGTAGATATGGTCAATTCCTTTGGCATTATACAA	1202
Db	1166	ACTAAACCATCTGTGAACAGTGTGGTTGTAAAGATATGGTCATTTCTTTTGGCATTTACATA	1225
QY	1203	GGATAAAATTTACATGTATGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGAGTT	1262
Db	1226	GGATAAAATTTACATGTATGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGAGCT	1285
QY	1263	GAGAGTTTTTCACATTCATATAGTCAATGGCTGTGTGTGACCCCTTAAGCAAAAGGACGA	1322
Db	1286	GAGAGTATTTCAATATTCATATGAATCATGGGTATTTGTAATCTCCGAAAGCTAAGATCA	1345
QY	1323	GTATGCAATGTTGGGCATCTCTGCACACATTTGTTTACACTGAAGAAATGGCCGAGTGTTCAT	1382
Db	1346	GTATGCAATGTTGGGCATCTCTGCACACATTTGTTTACACTGGCATCTGGCCCTGTGTGTTCAT	1405
QY	1383	GCTGCTCATCTTTGCTCACTGCCCTCTCTATGGATATATAGCAATCTGCAGGAATATGA	1442
Db	1406	GTTGGTCACTTTGCTCACTGCCCACTCTATGGATATATAGCGTTGTGCAGGAATATGA	1465
QY	1443	TTTGGATAAGAACACATGGAGTATTTACACACCCAGGGTGCCTTGTGCAAGGGGGTTA	1502
Db	1466	CTTGGAAAAAACACATGGAGTATTTACATCTACAGGGTGTCTTGTGCAAGGGGGTTA	1525
QY	1503	CGGCCATGACAGTGTTCACGACATAGGACCGAGGGCCCTATAGTTTCATCGGTGGCTACAA	1562
Db	1526	TGGCCACAGTAGTGCCTTATGATGACGAGCAAGAGGCTCTGTAGCTTCATCGTGGCTACAA	1585
QY	1563	GGCTTTTCAGTGGCCAAATAGTACCGGGCTTGAGATGATCTCTACCGATATGATGTGGATAC	1622

Db	1586		GGCTTTTCAGCGCCAAACAAATACCGGCTTGAGATGACCTCTACAGATACGATGTGGATAC	1645
Qy	1623		CCAGATGTGAGACCAATTCCTTAAGAGACGCCGCAATTTTCGGTTACTTCGACACAGCTGTGAT	1682
Db	1646		TCAGATGTGGACCAATTCCTTAAGAGACGCCGCAATTTTCGGTTACTTCGACACAGCTGTGAT	1705
Qy	1683		AGTCAGTGGAAACCAATGCTGCTGTTGGGGGAAAACACACAATGACACATCTATGAGCCA	1742
Db	1706		AGTGAGTGGAAACCAATGCTGCTGTTGGAGGGAACACACAATGACACATCTATGAGCCA	1765
Qy	1743		TGGGCCCCAAATAGCTTCTCTTCACAGATTCATGGGCTATGACATGTCCTGAGACCGCTGGTTC	1802
Db	1766		CGGTGCCAAATAGCTTCTCTTCACAGATTCATGGGCTATGACATGTCCTGAGACCGATGGTTC	1825
Qy	1803		AGTCTCTCCAGACCTTGATCTCCACCATGATGTCACACAGATTTGGCCATTTACAGACGTCTTT	1862
Db	1826		AGTCTCTCCAGACCTTGATCTCCACCATGATGTCACACAGATTTGGCCATTTACAGACGTCTTT	1885
Qy	1863		ACACAACAGCAACCATGATGATGTTTCGGTGGTTTCAATAGTTCCTCTCCTCAGGACATCTCT	1922
Db	1886		GTACAAACAGACCAATGATGTTTCGGCGGCTTCAACAGCCTCCTCCTCAGTGAACGCTCTT	1945
Qy	1923		GCTATTCAACCTCGGAAACAGTGTGATCGGATCGGAGTGAAGCGCTGTTTACGACGACG	1982
Db	1946		GGTCTTTTACCTCGAGCGAGTGCATGCAACACGACGTGAAGCTGTTGTGTGCGACGAGG	2005
Qy	1983		ACCTGGTATTTCGGTGTGTGTGGAAACACAGGGTTCGCTCAGTGTATCTCGTGGGCGCTGGC	2042
Db	2006		ACCTGGTATTTCGGTGTGTGTGGAAACACAGTTCGCTCGATGTACCTCTCGGAGTTGGC	2065
Qy	2043		AACGTATGACACAGAAAGATTAATTCAGAAATGTTTTTCCAAAGAACACTCTTGACCA	2102
Db	2066		AACCTGAAGAACACAGCAGAAAGTTAAATTCAGAGTCTTTTTCTAAAGAAACCTTGACCA	2125
Qy	2103		TGACAGATGTGACCCAGCACACACAGATTTGTTCACAGCTGTACAGGCCAACACCAATGAC	2162
Db	2126		TGACAGATGTGACCCAGCACACACAGATTTGTTCACAGCTGTACAGGCCAACACCAATGAC	2185
Qy	2163		CTGGTGAATGACCAATGTGTCCCGAGGAAACACACAGCTGCTCAGAAAGGCAGATCTCCAT	2222
Db	2186		CTGGTGAATGATCACTGTGTCCCTGTGAACACACAGCTGCACAGAAAGGCCAGATCTCCAT	2245
Qy	2223		TTTTAGGTATGAAATGCCCCAAGGATAACCCCATGTACTACTGTAAACAAGAACCCAG	2282
Db	2246		TGCCAAGATGAGAGTTGCCCAAGGATAACCCCATGTACTACTGCATTAAGAAACCAG	2305
Qy	2283		CTGCGAGAGTGTGCCCTGGACAGAACTGCCAGTGGGAGGCCCGGGAATCAGAGTGTGAT	2342
Db	2306		CTGCGAGAGTGTGCCCTGAGACAGAACTGCCAGTGGGAGGCCCGGGAATCAGAGTGTGAT	2365
Qy	2343		TGCGCTGCCCGAAAATCTGTGGCAATGGCTGGCAATTTGGTTGGAACCTCATGTTTGA	2402
Db	2366		CGCGCTGCCCGAAAATCTGTGGCAATGGCTGGCAATTTGGTTGGAACCTCATGTTTGA	2425
Qy	2403		AATTACTACTGCCAAGGAAATATGACAAATGCTAAATTTGTTCTGTAGGAAACCAATGC	2462
Db	2426		AATCACTACTGTCTAAGSAGAATATGACAAATGCTAAATTTGCTCCTGTAGGAACCAATGC	2485
Qy	2463		CTTTTGGCTTCTCTTACAAACCCAGAAAGGTAGAAATTTGCTTAAAGCAGCTCCGAAT	2522
Db	2486		CTTTTGGCTTCCCTCACATCCAGAGNAGGTGGAGCTTGCTTAAAGCAGCTTCGATT	2545
Qy	2523		AATGCAGTCACTCTCAGAGCAATGTCAGAGTCACTACCTTAAACCCCATGGGTGGCTTCGAA	2582
Db	2546		AATGCAATCACTCTAAAGTATGTCAAAGCTCACTCTGACTCCATCGGTTGGTCTTCGAA	2605
Qy	2583		GATCAATGTCTCTACTGGTGCTGGGAGATATGTCGCCATTTACAAATAGTTTACTACA	2642
Db	2606		GATCAATGTCTTACTGGTGCTGGGAGATATGTCGATTCACAAATAGTTTCTGCTGCA	2665
Qy	2643		GTGGATGCCGTCTGAGCCCGAGTGATCGGATTCGTGGAAATTTTATCAGAAACCCAGTAC	2702





Db 673 GAGAACTAACTGGCTCTCTCGATTGTGTAAACAGATGGACCTGGGAATTTAATAATAAAGA 732  
Qy 467 CGAAGTGACGTGGCTCATTTGAAGACACGCCAAATAGAAATAGAGACTTGGTTCAATC 526  
Db 733 CGAAGTGACATGGCTCATTTGAAGACACGCCAAATAGAAATAGAGACTTGGTTCAATC 792  
Qy 527 ATTITGCTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATG 586  
Db 793 ATTITGCTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATG 852  
Qy 587 CACCGTAGTGTCTGATTTAGTGCCCTCATTTGTTCCCTGAGAGAGATGGCAATGAGACTG 646  
Db 853 CACCTCTGATTTGCTGCTTCTAGTGCCCTCATTTGTTCCCTGAAAGAGATGGCAATGAGACG 912  
Qy 647 TCCCTGAGGTTGTTGCCACATCAGGTTATGCTTCTGCTGCATTTTTTTAGTGCATGCTGTT 706  
Db 913 CTCTGAGGTCAGTCTCACTTCAGGTTATGACTGCTGCATTTTTTTAGTGCATGCTGTT 972  
Qy 707 ATAAITGAGTGGATTTAATATTAATTAAGTTTATGATGTTTATGATGTTTATGATGTTTATG 766  
Db 973 ATAAITGAGTGGATTTAATATTAATTAAGTTTATGATGTTTATGATGTTTATGATGTTTATG 1032  
Qy 767 GCCGAGGAGGTGAAGATCAGTATAGCAGGAACTGTTGAATGTGAATGTTCTGAA 826  
Db 1033 GCCGAGGAGGTGAAGATCAGTATAGCAGGAGGCTGTTGAGTGTGAATGTTCTGAA 1092  
Qy 827 ACTGAAAGGTGAAGATCAGTATAGCAGGAACTGTTGAATGTGAATGTTCTGAA 886  
Db 1093 ACTGAAAGGTGAAGATCAGTATAGCAGGAACTGTTGAATGTGAATGTTCTGAA 1152  
Qy 887 GAGGATCTGCATTTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 946  
Db 1153 GAGGATCTGCATTTCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1212  
Qy 947 CTGGATGTTCTGCTTCTGCTACAGCTAACAGTCAATTTGACCTCGAGAGGAATATTCTA 1006  
Db 1213 CTGGATGTTCTGCTTCTGCTACAGCTAACAGTCAATTTGACCTCGAGAGGAATATTCTA 1272  
Qy 1007 ACTTAAAGCTCCAGAGATCTCATAAAGCTGTGTCATATGAGGAAATATTGAGGTTG 1066  
Db 1273 ATTAAAGCTCCAGAGATCTCATAAAGCTGTGTCATATGAGGAAATATTGAGGTTG 1332  
Qy 1067 TTGAGGATATATGTTCAACACTCAGATTAATACATGTTTCTAGCGTATGACCTTCTCT 1126  
Db 1333 TTGCGGATATATGTTCAACACTCAGATTAATACATGTTTCTAGCGTATGACCTTCTCT 1392  
Qy 1127 CTAGGAGTGTCTTCCACTAAACCGTTCTGTGAACATTTGTTGTTTATGATATGTTGAT 1186  
Db 1393 CTAGGAGTGTCTTCCACTAAACCGTTCTGTGAACATTTGTTGTTTATGATATGTTGAT 1452  
Qy 1187 CTTTGGCATTTATCAAGATTAATTTACATGTTATGAGGAAATTTGATCCACTGGA 1246  
Db 1453 CTTTGGCATTTATCAAGATTAATTTACATGTTATGAGGAAATTTGATCCACTGGA 1512  
Qy 1247 ATGTGACCAATGAGTTGAGAGTGTCTTCACTTCAATATGAGTCAATGAGTGTGTTGACCC 1306  
Db 1513 ACGTGACCAATGAGTTGAGAGTGTCTTCACTTCAATATGAGTCAATGAGTGTGTTGACCC 1572  
Qy 1307 CTAGGCAAGAGGAGTATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1366  
Db 1573 CGAAGCTAAGGATCAGTATGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1632  
Qy 1367 ATGCGCGAGTGTCTGCTGCTCATTTTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1426  
Db 1633 CTGCGCGTGTGTCTGCTGCTCATTTTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1692  
Qy 1427 ATGTGCGGAATATGATTGGATTAAGAACATAGGATATATTAACACCCAGGGTGCCT 1486  
Db 1693 TTGTGCGAGGAATATGATTGGATTAAGAACATAGGATATATTAACACCCAGGGTGCCT 1752  
Qy 1487 TTGTGCGAGGGGTTTACGGCCATAGCAGTGTGTTTACGACCATAGGACCCAGGGCTTATAG 1546  
Db 1753 TTGTGCGAGGGGTTTACGGCCATAGCAGTGTGTTTATGATGACCAACAGGCTCTGTAGC 1812

Qy 1547 TTCTATGTTGCTACAAAGGCTTTTCAGTGTCCAAATAGTACCGGCTTGAGATGATCTCTACC 1606  
Db 1813 TTCTATGTTGCTACAAAGGCTTTTCAGTGTCCAAATAGTACCGGCTTGAGATGATCTCTACC 1872  
Qy 1607 GATATGATGTTGATACCCAGATGTGACCAATCTTTAAGACAGCCGATTTTCCGTTACT 1666  
Db 1873 GATACGATGTTGATACCCAGATGTGACCAATCTTTAAGACAGCCGATTTTCCGTTACT 1932  
Qy 1667 TGACACAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1726  
Db 1933 TGACACAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1992  
Qy 1727 ACACATCTATGAGCCATGCGCAATGCTTCTTTCAGATTTTCATGCGCTTATGATGATG 1786  
Db 1993 ACATCTTCCATGAGCCATGCGCAATGCTTCTTTCAGATTTTCATGCGCTTATGATGATG 2052  
Qy 1787 CTGTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1846  
Db 2053 CTGTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2112  
Qy 1847 GCCATTCAGCAGTCTTACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1906  
Db 2113 GCCATTCAGCAGTCTTACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2172  
Qy 1907 TCCTCAGCAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1966  
Db 2173 TCCTCAGCAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2232  
Qy 1967 CTTGTTTACAGCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2026  
Db 2233 CTTGTTTACAGCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2292  
Qy 2027 TCTCTGCGGCTGCGCACTGATGAACAGAAAGAAAGTTAAATCAGAACTGTTTTTCA 2086  
Db 2293 TCTCTGCGGCTGCGCACTGATGAACAGAAAGAAAGTTAAATCAGAACTGTTTTTCA 2352  
Qy 2087 AAGAACTCTTGACCATGACATGTGACACAGACACAGATTTTACAGCTGTACAGCA 2146  
Db 2353 AAGAACTCTTGACCATGACATGTGACACAGACACAGATTTTACAGCTGTACAGCA 2412  
Qy 2147 ACACCAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2206  
Db 2413 ATACCAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2472  
Qy 2207 AAGGCGCAGTCTCCATTTTATGATGAAATTTGCTGCAATGAGTGTGCTGCAATGAGTGT 2266  
Db 2473 AAGGCGCAGTCTCCATTTTATGATGAAATTTGCTGCAATGAGTGTGCTGCAATGAGTGT 2532  
Qy 2267 GTAAACAGAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2326  
Db 2533 GTAAACAGAGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2592  
Qy 2327 GGAATCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2386  
Db 2593 GGAATCAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2652  
Qy 2387 GAACTCATGTTTGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2446  
Db 2653 GAACTCATGTTTGAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2712  
Qy 2447 GTAGCAACACAAATGCTTCTTTCAGTGTGCTTTCAGTGTGCTTTCAGTGTGCTTTCAGT 2506  
Db 2713 GTAGCAACACAAATGCTTCTTTCAGTGTGCTTTCAGTGTGCTTTCAGTGTGCTTTCAGT 2772  
Qy 2507 TTAAGCAGCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2566  
Db 2773 TTAAGCAGCTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2832  
Qy 2567 GGGTCTGCTGCGGAGATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2626  
Db 2833 GGGTCTGCTGCGGAGATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2892



```
QY 2627 CAATAGTTTACTACAGTGGATCGGCTGAGCCAGTGTGATGCTGGATTCTGTGGAATTT 2686
Db 2893 CAATAGTTTCTGAGTGGATGCACTGAGCCAGTGTGATGCTGGCTTCTGTGGATCT 2952
QY 2687 TATCAGAACCCAGTACTCGGGAATGAAGGCTGCAACCTGTCATCAACCCACTCAATGTA 2746
Db 2953 TGTGAGACCTAGTACTCGGGAATTAAGGCTGCAACCTGTCATCAACCCCTCTCAATGGCA 3012
QY 2747 GTGCTGTGAAAGGCTGCAACCCAGTGTGATGCTGAGTGGCCGACACCATGTGCTTGA 2806
Db 3013 GCGTGTGTGAAAGGCTGCAACCCAGTGTGATGCTGAGTGGCCGACACCATGTGCTTGC 3072
QY 2807 GGACAGCATGTGGAGATGCAACCCAGTGTGATGCTGAGTGGCCGACACCATGTGCTTGC 2866
Db 3073 GGACAGCATGTGGAGATGCAACCCAGTGTGATGCTGAGTGGCCGACACCATGTGCTTGC 3132
QY 2867 AGCAGTGTGGAATCCAAATGCTATGCTGCTTCTTCCCTTTTGGCCAGTGTGATGAAT 2926
Db 3133 AGCAGTGTGGAATCCAAATGCTATGCTGCTTCTTCCCTTTTGGCCAGTGTGATGAAT 3192
QY 2927 GTTATACATGAGCACCTGCCCTGAAATTTGTTAGGCTACTGTACCTGTAGTCAAT 2986
Db 3193 GGTATACATGAGCACCTGCCCTGAAATTTGTTAGGCTACTGTACCTGTAGTCAAT 3252
QY 2987 GCTTGGACCAACAGGCTGTGCTGGTACTGATCCAGCAATCTGCGCAAGGGAAT 3046
Db 3253 GCTTGGACCAACAGGCTGTGCTGGTACTGATCCCTAGCAATCTGCGCAAGGGAAT 3312
QY 3047 GATAGAGGTTCTTAAAGGACAGTGAAGTCTTCCAGGCTTCCAGGCTTCCAGGAAAT 3106
Db 3313 GTATTGAGGCGAGTATAAGGACCTGTAAGATCCCTGAGGCTTCTGAGGAAATG 3372
QY 3107 TCTATCCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGACAGCAGATACAACTGGT 3166
Db 3373 TGTATCCACAGCCCTTCTGAACTCCAGCATGTGTCTAGAGACAGCAGATACAACTGGT 3432
QY 3167 CTTTCAATTCAGTGTCCAAATGCAACGGCCAGCATGTAATGTCATCATCAGCA 3226
Db 3433 CTTTCAATTCAGTGTCCAAATGCAACGGCCAGCATGTAATGTCATCATCAGCA 3492
QY 3227 TCTGTGAGAGTGTGAGACCTGACACAGGCAAGCTTCCAGGCTTCCAGGCTTCCAGG 3286
Db 3493 TCTGTGAGAGTGTGAGACCTGACACAGGCAAGCTTCCAGGCTTCCAGGCTTCCAGG 3552
QY 3287 TCTAGCGTATCCCAATGAGGGAATGTGAGCATGTGAGTGAATGAGGCAAGCGT 3346
Db 3553 TCTATGTTGACCCGACTAATGAGGCAAGTGTGAGCATGTGAGTGAATGAGGCAAGCG 3612
QY 3347 CTCTGTGCAACCAACAGGCAAGTGTCTTCCAGGCAAGCTTCCAGGCTTCCAGG 3406
Db 3613 CACTGTGCAACCAACAGGCAAGTGTCTTCCAGGCAAGCTTCCAGGCTTCCAGG 3672
QY 3407 AGTGCAGCTATGTGAGGTGAGAAATCGATACCAAGGAACCTCTCAGAGGAACATGTT 3466
Db 3673 AGTGCAGCTATGTGAGGTGAGAAATCGATACCAAGGAACCTCTCAGAGGAACATGTT 3732
QY 3467 ATTATCTCTTCTATGATCATGATGATGATGATGATGATGATGATGATGATGATGAT 3526
Db 3733 ACTATACCTTCTCATTGATCATGATGATGATGATGATGATGATGATGATGATGATG 3792
QY 3527 ATTACACAGCTATCAATTTTGTGGCTTCTTCTGAGCAACCAAGGATTTGGAATGT 3586
Db 3793 ACTACACAGCTATCAATTTTGTGGCTTCTTCTGAGCAACCAAGGATTTGGAATGT 3852
QY 3587 TCATCAATGCTTCCAAAGATTTCAACCTCAACATCACTGGGCTGCGAGTTTCTCAGCTG 3646
Db 3853 TCATCAATGCTTCCAAAGATTTCAACCTCAACATCACTGGGCTGCGAGTTTCCAGCGG 3912
QY 3647 GAACCCAGGCTGAGAGAGATGCTGTTTCTCAAAACCAACATTAAGGATGATCAAG 3706
Db 3913 GAACCCAGGCTGAGAGAGATGCTGTTTCTCAAAACCAACATTAAGGATGATCAAG 3972
QY 3707 ATAGTTTCTCTAATGAGAAGTTTGAATTTTCGCAACCAACCAACCAATATCACTTTT 3766
```

```
Db 3973 ATAGCTTCTCTAATGAGAATTTGAATTTTCGCAACCAACCAACCAATCACTTTCTTGT 4032
QY 3767 ATGTCAGTAAATTTCACTGGCCATCAAAATTCAGGT 3803
Db 4033 ATGTCAGTAAATTTCACTGGCCATCAAAATTCAGAT 4069

RESULT 12
LOCUS AB093245 5683 bp mRNA linear ROD 07-FEB-2003
DEFINITION Mus musculus mRNA for mKIAA0548 protein.
ACCESSION AB093245
VERSION AB093245.1 GI:26006172
KEYWORDS FLI CDNA.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 Okazaki,N., Kikuno,R., Ohara,R., Inamoto,S., Hara,Y., Nagase,T.,
Ohara,O. and Koga,H.
Prediction of the coding sequences of mouse homologues of KIAA
gene: I. The complete nucleotide sequences of 100 mouse
KIAA-homologous cDNAs identified by screening of terminal sequences
of cDNA clones randomly sampled from size-fractionated libraries
Unpublished
2 (bases 1 to 5683)
Okazaki,N., Kikuno,R., Nagase,T., Ohara,O. and Koga,H.
Direct Submission
Submitted (04-OCT-2002) Hisashi Koga, Kazusa DNA Research
Institute, Laboratory for Genome Informatics; 2-6-7
Kazusa-kamata-ri, Kisarazu, Chiba 292-0818, Japan
(E-mail:mouse@kazusa.or.jp, Tel:81-438-52-3919, Fax:81-438-52-3918)
The CREATE program supported by Japan science and technology
corporation; cDNA full insert sequencing; Kazusa DNA Research
Institute; cDNA library construction, clone selection and 5'- &
3'-end one pass sequencing.
Location/Qualifiers
1..5683
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="mbg03041"
/tissue_type="brain"
/dev_stage="adult"
/notes="vector:modified pBC SK-"
1..5683
/gene="mKIAA0548"
<1..3181
/gene="mKIAA0548"
notes="CDS is predicted by in silico analysis. Start codon
is not identified."
/codon_start=2
/evidence="not experimental"
/product="mKIAA0548 protein"
/db_xref="GI:26006173"
/translation="AYDLISRWLPLNHSVSVVRYGSHLSLHKDKIYVYGGKIDST
GNVNEURPHIHNSWLLTPKAKQYAVGHSALHTLASGVNVLVIFGCPYLG
YISVQBYDLEKNTWSILHTGALVGGYHSSVYDDRTKALIVHGKIKAFSANKVRL
ADLYRYDVTQMTWKLDKSRFFRYLTVIVSGTVMVFGNTNLLSMGSAKCPSS
DFMAYDIACDRWSVLPRLHVDNFRFHSALVYNSMTYVFGFNLLSLDLVLTLE
QCDHRSEACAAVAGPGRICLDTQSRYCTSWELATBEQAEKLSKESKRTLDHRC
DQHTDCYCTANTNDCHWNCDCVPVNSCTEGQISIAKYESCPKDNPMYCNCKTSC
RSCALDNQCWEPNRQECIALPENI CNGWHLYGNSCLKITAKENYDNKLSRNNH
AFIASLTSQKKVFLVQLRMQSSQSKSLTTPWGLKINVSCHWEDMSPFFNS
LLOMPSPEDAGFCGILSEFTRGLKAATCINPLSGSVCTERFANHSKACQRTPCALR
TACGECTSSSECCMWCNSMKQCVDSNAYVASFPQCMVEYTWSSCPNCSGYTCS
HCLFOPCGCTDPTDNGKIGESYKGVKMPQSAGNVPQPLNLSMCLPSR
YNSWF IHPACQCNKSHSKINOSVCEKEDLTGKHCTCISGFGYDPTNGKQCPCK
CNGHASLCTNTGKCFCTTKYKGDQCLCEVENRYGNPLKGTTCVYTLIDYQTFPS
LSQDDRRYTAINFATVPDEQNRLDNFINASKNFNLNITWATSPFAGTGTGEVPVW
```

SKTTIKYKDSFNSNEKFDPRNHNITFFVVSNTFWPIKIQIAFSQHSNFMDLVQFFV  
TFPSCPLSLLVAAVVKIKQCSWASRRREQLLRMOQMASRPPASVNVVALETDREPP  
DLIGSISKVPEKPIALEPCFNGKAALVLSVFLPRGLGGIIPPGSGGLAVASALVDIS  
QOMPVYKESGAVNRKQOPPAQPGTCT

ORIGIN

Query Match		59.6%	Score 2276.4;	DB 10;	Length 5683;
Best Local Similarity		90.3%;	Pred. No. 0;		
Matches 2433;		Conservative 0;	Mismatches 261;	Indels 0;	Gaps 0;
QY	1110	ACGGTATGACCTTGTCTTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGGT	1169		
DB	1	ACGGTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCAATCTGTGAACAGTGTGGT	60		
QY	1170	TGTTAGATATGGTCATTCTTTGGCATTATACAGGATMAAATTTACATGTATGGAGGAAA	1229		
DB	61	TGTAAGATATGTCATTCTTTGGCATTATACAGGATMAAATTTACATGTATGGAGGAAA	120		
QY	1230	AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCATTCATTAATGATC	1289		
DB	121	AATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCATTCATTAATGATC	180		
QY	1290	ATGGGTGTTGTTGACCCCTAAGCGAAAGGAGCAGATGAGTGGTGGGCACTCTGCACA	1349		
DB	181	ATGGGTGTTGTTGACCCCTAAGCGAAAGGAGCAGATGAGTGGTGGGCACTCTGCACA	240		
QY	1350	CAATTGTTACATGGAAGAAATGGCCGAGTGGTCAATGCTGGTTCATTTTGGTCACTGCCCTCT	1409		
DB	241	CAATTGTTACATGGAAGAAATGGCCGAGTGGTCAATGCTGGTTCATTTTGGTCACTGCCCTCT	300		
QY	1410	CTATGGATATAGCAATGTCAGGAATATGATTTGGATAGGACATGAGTATATT	1469		
DB	301	CTATGGATATAGCAATGTCAGGAATATGATTTGGATAGGACATGAGTATATT	360		
QY	1470	ACACACCCAGGGTGCCCTTGTCAAGAGGGGGTTACCGGCATAGCAGTGTTTACGACCATTAG	1529		
DB	361	ACATACTCAGGGTGTCTTTGTGCAAGGGGGTTATGCCCCACAGTAGTGTATATGACAG	420		
QY	1530	GACACGGGCCCTATAGCTTCATGTTGGCTACAGGCTTTAGTGGCCATAGTACCGGCT	1589		
DB	421	GACCAAGGCTCTGTACGTTTCATGTTGGTGTACAGGCTTTACGCGCCACAAATACCGGCT	480		
QY	1590	TGCAGATGATCTCTACCGATATGATGTGATACCGATGCGACCATTTCTTTAAGGACAG	1649		
DB	481	TGCAGATGATCTCTACCGATATGATGTGATACCGATGCGACCATTTCTTTAAGGACAG	540		
QY	1650	CGATTTTCCGTTACTTGCACACAGCTGTGATAGTGGAGAACCAATGCTGGTGTGG	1709		
DB	541	CGATTTTCCGTTACTTGCACACAGCTGTGATAGTGGAGAACCAATGCTGGTGTGG	600		
QY	1710	GGGAACACACACATGATCTATGAGCCATGCGCCCAATGCTTCTCTTCAGATTT	1769		
DB	601	AGGGAACACACACATGATCTATGAGCCATGCGCCCAATGCTTCTCTTCAGATTT	660		
QY	1770	CATGCCCTATGACATGCTGTGACCGCTGTGATGCTTCCAGACCTGATCTCCACCA	1829		
DB	661	CATGCCCTATGACATGCTGTGACCGCTGTGATGCTTCCAGACCTGATCTCCATCA	720		
QY	1830	TGATGTCAACAGATTTGGCCATTAGCAGCTCTTACACAGCAGCAGTATGTTCCG	1889		
DB	721	TGATGTCAACAGATTTGGCCATTAGCAGCTCTTACACAGCAGCAGTATGTTCCG	780		
QY	1890	TGGTTTCAATAGTCTCTCTCAGCAGCATCTGATTTACCTTCGGAACAGTGTGATGC	1949		
DB	781	CGGCTTCAACAGCTCTCTCTCAGTACGCTTGTGTCTTTACCTTGGAGCAGTCCGATGC	840		
QY	1950	GCATCGGAGTGAAGCGCTTGTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAAACAC	2009		
DB	841	ACACCGCAGTGAAGCGCTTGTGGCAGCAGGACCTGGTATTCGGTGTGTGTGGGACAC	900		
QY	2010	AGGGTCTGCTCAGTGTATCTCTGGGGCGCTGGCACTGATGAACAGAAAGATTAAA	2069		
DB	901	ACAGTCTGCTCAGTGTATCTCTGGGGCGCTGGCACTGATGAACAGAAAGATTAAA	960		

QY	2070	ATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACACGACACAGATTG	2129		
DB	961	ATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACACGACACAGATTG	1020		
QY	2130	TTACAGCTGTACAGCCCAACCAATGACTGCCACTGGTGAATGAACATTTGTGCCCCAG	2189		
DB	1021	TTACAGCTGTACAGCCCAACCAATGACTGCCACTGGTGAATGAACATTTGTGCCCTGT	1080		
QY	2190	GAACCAAGCTGTCTCAGAGGCCAGATCTCCATTTTGTAGTATGAGAAATGCCCCAAGA	2249		
DB	1081	GAACCAAGCTGTCTCAGAGGCCAGATCTCCATTTTGTAGTATGAGAAATGCCCCAAGA	1140		
QY	2250	TAAACCCATGCTACTCTGTAAACAAGAACCACTGCGAGGCTGTGCCCTGACACAGAA	2309		
DB	1141	TAAACCCATGCTACTCTGTAAACAAGAACCACTGCGAGGCTGTGCCCTGACACAGAA	1200		
QY	2310	CTGCCAGTGGAGCCCGGGAATCAGAGTGCAATGCGCTGCCCGGAAATATCTGTGGCAT	2369		
DB	1201	CTGCCAGTGGAGCCCGGGAATCAGAGTGCAATGCGCTGCCCGGAAATATCTGTGGCAA	1260		
QY	2370	TGGCTGCACTTTGTTGGAAACTCATGTTTGAATAATTAATCTGCAAGAGAAATATGA	2429		
DB	1261	TGGCTGCACTTTGTTGGAAACTCATGTTTGAATAATTAATCTGCAAGAGAAATATGA	1320		
QY	2430	CAATGCTAAATTTCTGTAGGAACCAATGCGCTTTTGGCTTTCTTTACAAACCCAGAA	2489		
DB	1321	CAATGCTAAATTTCTGTAGGAACCAATGCGCTTTTGGCTTTCTTTACAAACCCAGAA	1380		
QY	2490	GAAGGTAGATTTGCTTAAAGCAGCTGCAATTAATGAGTCACTCTCAGAGCATGTCCAA	2549		
DB	1381	GAAGGTAGATTTGCTTAAAGCAGCTGCAATTAATGAGTCACTCTCAGAGCATGTCCAA	1440		
QY	2550	GCTCAGCTTAAACCCATGGCTGGCCCTCGGAAGATCAATGTGCTCTACTGTGCTGGGA	2609		
DB	1441	GCTCAGCTTAAACCCATGGCTGGCCCTCGGAAGATCAATGTGCTCTACTGTGCTGGGA	1500		
QY	2610	AGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCGCTGTGAGCCAGTATGC	2669		
DB	1501	GGATATGTCCTCCATTTACAAATAGTTTGTGCTGAGTGGATGCCATCTGAGCCAGTATGC	1560		
QY	2670	TGATTTCTGTGGAATTTTATCAGAACCCAGTACTGCGGACCTGAAAGGCTGCAACCTGCAT	2729		
DB	1561	TGCTTCTGTGGGATCTTGTGAGAGCTTACTTCTGGGATTTAAAGGCTGCAACCTGCAT	1620		
QY	2730	CAACCCATCAATGCTAGTGTCTGTGAAAGGCTGCAACCCAGTGTCTAAGCAGTGGCG	2789		
DB	1621	CAACCCCTCTCAATGGCAGCTCTGTGAAAGGCTGCAACCCAGTGTCTAAGCAGTGGCG	1680		
QY	2790	GACACCATGTCCTTGAGGACAGCATGTGGAGATTGCAACCGGCGACCTCTGAGTGCAT	2849		
DB	1681	GACACCATGTCCTTGAGGACAGCATGTGGAGATTGCAACCGGCGACCTCTGAGTGCAT	1740		
QY	2850	GTGGTGCAGCAACATGAAGCAGTGTGTGACTTCCAAATGCTATGTGGCTCTCTCCCTTT	2909		
DB	1741	GTGGTGCAGCAACATGAAGCAGTGTGTGACTTCCAAATGCTATGTGGCTCTCTCCCTTT	1800		
QY	2910	TGGCCAGTGTATGGAAATGGTATACGATGAGCAGCTTCCCGCCCTGAAATTTGTTCAGGCTA	2969		
DB	1801	TGGCCAGTGTATGGAAATGGTATACGATGAGCAGCTTCCCGCCCTGAAATTTGTTCAGGCTA	1860		
QY	2970	CTGTACTCTGTAGTCAATTTGCTTGGAGCAACAGGCTGTGGCTGGTGTACTGTATGCCAGAA	3029		
DB	1861	TTGTACTCTGTAGTCAATTTGCTTGGAGCAACAGGCTGTGGCTGGTGTACTGTATGCCAGAA	1920		
QY	3030	TACTGGCAAGGAAATGCAATAGAGGTTCTTAAAGGAAACAGTGAAGATGCTTCCGCA	3089		
DB	1921	TACTGGCAAGGAAATGCAATAGAGGTTCTTAAAGGAAACAGTGAAGATGCTTCCGCA	1980		
QY	3090	AGCCCTCTACAGGAAATTTCTATCCACAGCCCTCTCAATTCACAGCATGTGTCTAGAGGA	3149		
DB	1981	AGCCCTCTACAGGAAATTTCTATCCACAGCCCTCTCAATTCACAGCATGTGTCTAGAGGA	2040		



1143 ACTAAACCGTTCTGTGAAACAATGTGGTGTGTAGATATGTCATCTTTGGCAATTATACAA 1202  
1140 ACTAAACCGTTCTGTGAAACAATGTGGTGTGTAGATATGTCATCTTTGGCAATTATACAA 1199  
1203 GGATAAAATTTACATGTATGGAGGAAAATTCATCCAACTGGGAATGTGACCAATGAGTT 1262  
1200 GGATAAAATTTACATGTATGGAGGAAAATTCATCCAACTGGGAATGTGACCAATGAGTT 1259  
1263 GAGAGTTTTCACATTCATAATGAGTCATGGGTGTGTGAGCCCTTAAGGCAAGGACCA 1322  
1260 GAGAGTTTTCACATTCATAATGAGTCATGGGTGTGTGAGCCCTTAAGGCAAGGACCA 1319  
1323 GTATGCAAGTGTGGGCACTCTGCACACATTCATACATGAGAAATGCGGAGTGGTCAT 1382  
1320 GTATGCAAGTGTGGGCACTCTGCACACATTCATACATGAGAAATGCGGAGTGGTCAT 1379  
1383 GCTGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGAGGAATATGA 1442  
1380 GCTGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGAGGAATATGA 1439  
1443 TTTGGTAAGAACACATGAGTATATACACACCCAGGTCCTTGTGCAAGGGGTTA 1502  
1440 TTTGGTAAGAACACATGAGTATATACACACCCAGGTCCTTGTGCAAGGGGTTA 1499  
1503 CGGCCATAGCAGTGTTCAGCACTAGGACCAAGGCGCTTATACGTTTCATGTGGCTACAA 1562  
1500 CGGCCATAGCAGTGTTCAGCACTAGGACCAAGGCGCTTATACGTTTCATGTGGCTACAA 1559  
1563 GCGTTTCAGTGCAATAGTACCGGCTTGCAGATGATCTTACCGATATGATGTGGATAC 1622  
1560 GCGTTTCAGTGCAATAGTACCGGCTTGCAGATGATCTTACCGATATGATGTGGATAC 1619  
1623 CCAGATGTGACCATCTTTAAGGACAGCGGATTTTCGTTTCTTCTGTCACACAGCTGTGAT 1682  
1620 CCAGATGTGACCATCTTTAAGGACAGCGGATTTTCGTTTCTTCTGTCACACAGCTGTGAT 1679  
1683 AGTGTGTGAAACCATGCTGGTGTGGGGGAAACACACACATGACATCTATGAGCCA 1742  
1680 AGTGTGTGAAACCATGCTGGTGTGGGGGAAACACACACATGACATCTATGAGCCA 1739  
1743 TGGCCCAAAATCTTCTTCAGATTTTCAGGCTTATGACATTCGCTGTGACCCCTGGTC 1802  
1740 TGGCCCAAAATCTTCTTCAGATTTTCAGGCTTATGACATTCGCTGTGACCCCTGGTC 1799  
1803 AGTGTCTCCGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1862  
1800 AGTGTCTCCGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGCAGTCTT 1859  
1863 ACACAGACCATGATGTGTTGGTGGTGTGATGATGATGATGATGATGATGATGATGATGAT 1922  
1860 ACACAGACCATGATGTGTTGGTGGTGTGATGATGATGATGATGATGATGATGATGATGAT 1919  
1923 GGTATTCACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCGCTTTTGTAGCAGCAGG 1982  
1920 GGTATTCACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCGCTTTTGTAGCAGCAGG 1979  
1983 ACCTGGTATTCGGTGTGTGGAACACAGGTCGTCTCAGTGATATCTGCTGGGCGCTGGC 2042  
1980 ACCTGGTATTCGGTGTGTGGAACACAGGTCGTCTCAGTGATATCTGCTGGGCGCTGGC 2039  
2043 AACTGATGAACAGAGAAAAGTTAAATCAGATGTTTTTCCAAAAGAACTCTTGACCA 2102  
2040 AACTGATGAACAGAGAAAAGTTAAATCAGATGTTTTTCCAAAAGAACTCTTGACCA 2099  
2103 TGACAGATGTGACCAAGCAACAGATGTTGATCAGCTGTACAGCCCAACCAATGATGCGCA 2162  
2100 TGACAGATGTGACCAAGCAACAGATGTTGATCAGCTGTACAGCCCAACCAATGATGCGCA 2159  
2163 CTGGTGCAATGACCATGTGTCGCCAGAACACAGCTGCTCAGAGCCAGATCTCCAT 2222  
2160 CTGGTGCAATGACCATGTGTCGCCAGAACACAGCTGCTCAGAGCCAGATCTCCAT 2219  
2223 TTTTAGTATGAGAATTTGCCCCCAAGGATAACCCCATGTACTGTGTAACAAGAGACCAAG 2282

2220 TTTTAGTATGAGAATTTGCCCCCAAGGATAACCCATATGTACTGTAAACAAGAGACCAAG 2279  
2283 CTGAGAGAGTGTGCTCCCTGGACCAAGAACTGCCAGTGGAGCCCGGAATCAGAGTGCAT 2342  
2280 CTGAGAGAGTGTGCTCCCTGGACCAAGAACTGCCAGTGGAGCCCGGAATCAGAGTGCAT 2339  
2343 TGCCCTGCCCGAAAATATCTGTGCAATGCTGCTGCAATTTGGTTGG 2387  
2340 TGCCCTGCCCGTAGGCTTGCAGGGTCATCTTGGTGTGTGGG 2384

RESULT 14  
ARI64813  
LOCUS ARI64813 2419 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 8 from patent US 6274339.  
ACCESSION ARI64813  
VERSION ARI64813.1 GI:16238022  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 2419)  
AUTHORS Moore, K. and Nagle, D. Lynn.  
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity  
JOURNAL Patent: US 6274339-A 8 14-AUG-2001;  
FEATURES  
Location/Qualifiers  
source 1..2419  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 39.8%; Score 1521.8; DB 6; Length 2419;  
Best Local Similarity 80.7%; Pred. No. 18-299;  
Matches 1895; Conservative 0; Mismatches 247; Indels 205; Gaps 2;  
3 GGTGGCGGACAGCGCGGCAACTGAGCAAGGCTGAGGAGGAGGAGCGCGGCGACCGCAGC 62  
82 GGTGGCGGCGCGCGGCGGCGACTGAGGCGGCTGAGGGGAGGAGCAAGAGACGACGAGC 141  
63 GCTGGCGGCGAGAGCGGCGGCGCGCACTGGAAGCTGGGACGTGACAGAGGCTGGAGGCG 122  
142 GCCTGGCGGCGAGAGGCGGCGGAGCAGCAGCCGCTGCAACCGGAGACAGGGGCGCTGAGG 201  
123 GGGGCTGGGCGGCGGCGCTGCGCTCCCGGCGCTGCTCTCCACCGCTGCGGCGCAGCGCT 182  
202 GGGACCGCGCGCGCGCTGCTCTCCCGGCGCTGCTCGGGGCGCTGCGGCGCGCTGCGGCG 261  
183 GCTGCTGCTGCTGCTGCTGCTCCCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 242  
262 GCT 312  
243 CGAGGCGGCGGCGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302  
313 CGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372  
303 TGACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362  
373 TGACCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432  
363 CGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410  
433 CACGGGCTGGTGGGCGAGCAATGCCAGCACTGCGGGGCGCGCTTTCAGGACATCTGCTC 492  
411 ----- 410  
493 AGCGCTATAATCACAGCTGTTCCGAGAGTGGAGTGGAGGAGCAAGTTCGAGGCGAGCTTC 552  
411 ----- 410  
553 GGCTACAGAAATAGTTTCAAGAGTAACCTGGGGCAACTTGGGCTTGTCTTCCAAAACCAAAA 612

411 ----- 410  
613 TGACGAAAGGACAGCTAGAGTCTTTTGGGAAATTTTAGCTGACTAAATTTTCAAC 672  
411 ----- ACTAATCTGATCTTCTGGGTTTGTGACAGATGACCTGGAATTTATAATACAAA 466  
673 GAGAACTAATCTGCTCTTCTGATTTGTAAACAGATGACCTGGAAATTTATAATATAAGA 732  
467 CGAAGTCACAGTGGCTCAATTAAGAGGACAGCCAAATAGAAATAATGAGACTTCGTTTCAATC 526  
733 CGAAGTGCATGCTCAATTAAGAGGACAGCCAAATAGAAATAATGAGACTTCGTTTCAAC 732  
527 ATTTTGTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATG 586  
793 ATTTTGTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTACG 852  
587 CACCGTAGTGTGCAATTTAGTGGCTCAATTTCTGAGAGAGATGGCAATGAGACTG 646  
853 CACCTCTGATGCTGCTTTAGTGGCTCAATTTCTGAGAGATGGCAATGAGAGCG 912  
647 TCCCTGAGGTTGTTGCCACATCAGGTTATGCTTGTGCAATTTTATGATGCTGCTT 706  
913 CTCCTGAGGTCACCTGCTCACTTCAGGTTATGCACTGCTGCAATTTTTCAGTGTGCTGCTT 972  
707 ATATTTGACTGGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 766  
973 ATATTTGACTGGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1032  
767 GCCGAGAGAGTGTAAAGTCAATAGTACAGCAAACTGTTGAAATGTGAATGTTCTGAAA 826  
1033 GCCGAGAGAGTGTAAAGTCAATAGTACAGCAAACTGTTGAAATGTGAATGTTCTGAAA 1092  
827 ACTGAAAGTGTAAAGTCAATAGTACAGCAAACTGTTGAAATGTGAATGTTCTGAAA 886  
1093 ACTGAAAGTGTAAAGTCAATAGTACAGCAAACTGTTGAAATGTGAATGTTCTGAAA 1152  
887 GAGCATCTGCAATTTCAAGTGTATGAGAGATGCTCTGCTTCTCAGATGCGCAGGCTC 946  
1153 GAGCATCTGCAATTTCAAGTGTATGAGAGATGCTCTGCTTCTCAGATGCGCAGGCTC 1212  
947 CTGATGTTGAGTCTCTACAGTAACTACAGTCAATTTGAGCTCGAGAGAAATTTCTA 1006  
1213 CTGATGTTGAGTCTCTACAGTAACTACAGTCAATTTGAGCTCGAGAGAAATTTCTG 1272  
1007 ACTTAAAGTCTCCAGAGATCTCAATAGCTGTGCTCAATGGAATTAATTTGAGTGG 1066  
1273 ATTTAAAGTCTCCAGAGATCTCAATAGCTGTGCTCAATGGAATTAATTTGAGTGG 1332  
1067 TTGAGGATATATGTTCAACCACTCAGATTTATAATGTTCTAGCTGATGACCTTGCTT 1126  
1333 TTGAGGATATATGTTCAACCACTCAGATTTATAATGTTCTAGCTGATGACCTTGCTT 1392  
1127 CTAGGAGTGGTCTTCACTAAACCGTCTGTGAACAATGTGGTGTAGATATGCTCAAT 1186  
1393 CTAGGAGTGGTCTTCACTAAACCGTCTGTGAACAATGTGGTGTAGATATGCTCAAT 1452  
1187 CTTTGGCATTATACAGGATAAAATTTACATGATGAGGAGAAATTTGATCAACTGGGA 1246  
1453 CTTTGGCATTATACAGGATAAAATTTACATGATGAGGAGAAATTTGATCAACTGGGA 1512  
1247 ATGTGACCAATGAGTGTAGAGTTTTCATATTCATATTAATGAGTCAATGTTGTGACCC 1306  
1513 ACGTGACCAATGAGTGTAGAGTTTTCATATTCATATTAATGAGTCAATGTTGTGACCC 1572  
1307 CTAAGCAAGAGATGATGAGTGTGGGCTCTGACACATTTGATCACTGGAAGA 1366  
1573 CGAAGCTAAGGATCAGTATGAGTGTGGGCTCTGACACATTTGATCACTGGAAGA 1632  
1367 ATGGCGGAGTGTGATGCTGCTGCTCATCTTTGGTCACTGCGCTCTCTATGATATATAGCA 1426  
1633 CTGGCGGTGTGCTGCTGCTGCTCATCTTGGTCACTTGGCTCACTCTATGATATATAGCG 1692  
1427 ATGTGAGGATATGATTTGGATTAAGAACACATGAGATATATACACCCAGGTTGCC 1486

1693 TTGTGCGAATATATGCTTTGGAAAGAACACATGAGATATATATCATCTAGGTGCTC 1752  
1487 TTGTGCAAGGGGGTTTACCGCCATAGCAGTGTTTTACGACCATAGGACCGGCTTATACG 1546  
1753 TTGTGCAAGGGGGTTTATGCGCCACAGTGTGTTTATGATGACAGGACCAAGGCTCTGTACG 1812  
1547 TTCTATGTTGGCTTAAAGGCTTTTCAAGTGTGCAATAGTACCGGCTTGCAGATGATCTCTACC 1606  
1813 TTCTATGTTGGCTTAAAGGCTTTTCAAGTGTGCAATAGTACCGGCTTGCAGATGATCTCTACA 1872  
1607 GATATGATGTGATACCCAGATGTGACCAATTTCTTAAAGGACAGCCGATTTTTCGCTTACT 1666  
1873 GATACGATGTGATACTCAGATGTGACCAATTTCTTAAAGGACAGCCGATTTTTCGCTTACT 1932  
1667 TCCACACAGCTGTGATGATGAGTGGAAACCATGCTGTTTTCGCGGAAACACACACAATG 1726  
1933 TGCATACAGCTGTGATGATGAGTGGAAACCATGCTGTTTTCGCGGAAACACACACAATG 1992  
1727 ACACATCTATGAGCCATGCGCCAAATGCTTCTTTCAGATTTTCATGGCTTATGACATTTG 1786  
1993 ACATTTCCATGAGCCATGCGCCAAATGCTTCTTTCAGATTTTCATGGCTTATGACATTTG 2052  
1787 CCTGTGACCGCTGCTCAGTGTCTTCCAGACCTGATCTCCACCATGATGTCAACAGATTTG 1846  
2053 CTTGTGACCGATGCTCAGTGTCTTCCAGACCTGATCTCCACCATGATGTCAACAGATTTG 2112  
1847 GCCATTCAGAGCTTTACACACAGCAGCATGTATGTTTTCGCTGTTTCAATAGTCTCTC 1906  
2113 GCCATTCAGAGCTTTTACACACAGCAGCATGTATGTTTTCGCTGTTTCAATAGTCTCTC 2172  
1907 TCTCTGAGCAGATCTCTGTTTATTCACCTCGAAGACAGTGTGATCGCATCGGAGTGAAGCG 1966  
2173 TCTCTGAGCAGCTCTGTTTATTCACCTCGAAGACAGTGTGATCGCATCGGAGTGAAGCG 2232  
1967 CTTGTTTACAGAGACCTGTTTATTCGCTGTTTTCGCTGTTTTCGCTGTTTCAATAGTCTCT 2026  
2233 CTTGTTGAGCAGCAGCAGCTGTTTATTCGCTGTTTTCGCTGTTTTCGCTGTTTCAATAGTCTCT 2292  
2027 TCTCTGAGCAGCTGTTTATTCACCTCGAAGACAGTGTGATCGCATCGGAGTGAAGCG 2086  
2293 CTTCTGAGGATTTGCAATGAGAAACAGCAGAAAGTTTAAATCAGATGTTTTCCTCA 2352  
2087 AAAGAACTCTTTCACCATGACAGATGTGACAGCAGACAGATTTTACAGTGTGACAGCA 2146  
2353 AAAGAACTCTTTCACCATGACAGATGTGACAGCAGACAGATTTTACAGTGTGACAGCA 2412  
2147 ACACCA 2153  
2413 ATACCA 2419

RESULT 15  
AR164815  
LOCUS AR164815  
DEFINITION Sequence 12 from patent US 6274339.  
ACCESSION AR164815  
VERSION AR164815.1 GI:16238024  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 6370)  
AUTHORS Moore, K. and Nagle, D. Lynn.  
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity  
JOURNAL Patent: US 6274339-A 12 14-AUG-2001,  
FEATURES Location/Qualifiers  
          1. 6370  
          /organism="unknown"  
          /mol\_type="unassigned DNA"

Query Match 35.7%; Score 1364.4; DB 6; Length 6370;  
Best Local Similarity 95.0%; Pred. No. 1.4e-267;  
Matches 1492; Conservative 7; Mismatches 48; Indels 23; Gaps 8;

QY 2257 ATGTACTACTTAACAAAGAACACAGCTGCAGAGCTGTGCCCTGGACAGAACTGCCAG 2316  
DB 1 ATGTACTACTTAACAAAGAACACAGCTGCAGAGCTGTGCCCTGGACAGAACTGCCAG 60  
QY 2317 TGGAGCCCGGAATCAGAGTGCAATGCCCTGCCGGAATAATCTGTGGCAATGGCTGG 2376  
DB 61 TGGAGCCCGGAATCAGAGTGCAATGCCCTGCCGGAATAATCTGTGGCAATGGCTGG 120  
QY 2377 CATTTGGTTGGAATCACTGTTTGAATAATTACTGCGCAAGGAGAAATTGACAAATGCT 2436  
DB 121 CATTTGGTTGGAATCACTGTTTGAATAATTACTGCGCAAGGAGAAATTGACAAATGCT 180  
QY 2437 AAATTGTTCTGTAGGAACCAATGCCCTTTTGGCTTTCTTTTACAAACCAAGAGGTA 2496  
DB 181 AAATTGTTCTGTAGGAACCAATGCCCTTTTGGCTTTCTTTTACAAACCAAGAGGTA 240  
QY 2497 GAATTGTCCTTAAGCAGCTGGAATAATGCAATCATCTCAGAGCATGTCGAAGCTCACC 2556  
DB 241 GAATTGTCCTTAAGCAGCTGGAATAATGCAATCATCTCAGAGCATGTCGAAGCTCACC 300  
QY 2557 TTAACCCCATGGTGGGCTT---CGGAAGATCAATGTGCTACTGG-----TGCTGG 2608  
DB 301 TTAACCCCATGGTGGGCTTTCGGGAAGTTCATGTGCTACTKGGTGTGGGAAG 360  
QY 2609 AGATATGTCCCATTTACAAATAG-TTACTACAGTGGATGCCGTCTG-----AGCCAG 2663  
DB 361 GATATGKTCCTTTTACAAATGTTTACTACAGTGGGATGCCGCTTTGAGGCCAG 420  
QY 2664 TGATGCT--GGATTCTGTGGAAATTTT-----ATCAGAACCCAGTACTCGGGGACTGAAG 2716  
DB 421 TGTTGCTTGGATTTCTGTGGAAATTTTATCAGAACCCAGTACTTCGGGGACTGAAG 480  
QY 2717 CTGCAACTGCAAT--CAACCCACTCAATGTGATGTCTGTGAAAGGCTGCAAAACCAAG 2774  
DB 481 CTGCAACTGCAATTCACCCACTCAATGTGATGTCTGTGAAAGGCTGCAAAACCAAG 540  
QY 2775 TGCTAA-GCAGTGGCGGACACCATGTGCTTTCAGACAGCATGTGGAGATGTCACACAG 2833  
DB 541 TGCTAAGGAGTGGCGGACACCATGTGCTTTCAGACAGCATGTGGAGATGTCACACAG 600  
QY 2834 GCAGCTCTGAGTGCAATGTGGTGCAGCAACATGAAGCAGTGTGGACTCCAATGCCATG 2893  
DB 601 GCAGCTCTGAGTGCAATGTGGTGCAGCAACATGAAGCAGTGTGGACTCCAATGCCATG 660  
QY 2894 TGGCTCTCTCCCTTTTGGCCAGTGTATGGATGGTATACATGACGACCTGCCCCCTG 2953  
DB 661 TGGCTCTCTCCCTTTTGGCCAGTGTATGGATGGTATACATGACGACCTGCCCCCTG 720  
QY 2954 AAAATTGTTCAAGCTACTGTACTGTACTGTCTTGGAGAACACAGGCTGTGGCTGGT 3013  
DB 721 AAAATTGTTCAAGCTACTGTACTGTACTGTCTTGGAGAACACAGGCTGTGGCTGGT 780  
QY 3014 GTACTGATCCAGCAATFACTGGGAAGGAATGATAGAGGTTCTCTAAGAGGACAG 3073  
DB 781 GTACTGATCCAGCAATFACTGGGAAGGAATGATAGAGGTTCTCTAAGAGGACAG 840  
QY 3074 TGAAGATGCTTCCGAGCCCTCAGGAAATTTCTATCCACAGCCCTGCTCAATTTCCA 3133  
DB 841 TGAAGATGCTTCCGAGCCCTCAGGAAATTTCTATCCACAGCCCTGCTCAATTTCCA 900  
QY 3134 GCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTTCAATCTGCTCCAGCTTGCCAA 3193  
DB 901 GCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTTCAATCTGCTCCAGCTTGCCAA 960  
QY 3194 GCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAACTGACCA 3253  
DB 961 GCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAACTGACCA 1020  
QY 3254 CAGCAAGCACTGCGAGACCTGCAATATCTGGCTTTCTACGGTGTATCCCAATGGAGGA 3313

DB 1021 CAGCAAGCACTGCGAGACCTGCAATATCTGGCTTCTACGGTGTATCCCAATATGGAGGA 1080  
QY 3314 AATGTCAAGCATGCAAGTGCATGGGACGCGTCTGTGTCAACACCAACACGGGCAAGT 3373  
DB 1081 AATGTCAAGCATGCAAGTGCATGGGACGCGTCTGTGTCAACACCAACACGGGCAAGT 1140  
QY 3374 GCTTCTGCACACCAAGGGGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAAATC 3433  
DB 1141 GCTTCTGCACACCAAGGGGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAAATC 1200  
QY 3434 GATACCAAGGAAACCTCTCAGAGGACATGTTTATATATCTCTTATTTGACTATCAGT 3493  
DB 1201 GATACCAAGGAAACCTCTCAGAGGACATGTTTATATATCTCTTATTTGACTATCAGT 1260  
QY 3494 TCACCTTTAGTCTATCCAGGAAGATGATGGTATTTACACAGCTATCAATTTTGGGCTA 3553  
DB 1261 TCACCTTTAGTCTATCCAGGAAGATGATGGTATTTACACAGCTATCAATTTTGGGCTA 1320  
QY 3554 CTCTGACGAAACCAACAGGATTTGGACATGTTTCAATGCTCCAGAAATTTCAACC 3613  
DB 1321 CTCTGACGAAACCAACAGGATTTGGACATGTTTCAATGCTCCAGAAATTTCAACC 1380  
QY 3614 TCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCCCTG 3673  
DB 1381 TCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCCCTG 1440  
QY 3674 TTGTTTCAAAACCAACATTAAGGAGTACAAAGATGTTTCTCTAATGAGAAATTTGATT 3733  
DB 1441 TTGTTTCAAAACCAACATTAAGGAGTACAAAGATGTTTCTCTAATGAGAAATTTGATT 1500  
QY 3734 TTGCGAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAAATTTCACTGCGCCATCA 3793  
DB 1501 TTGCGAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAAATTTCACTGCGCCATCA 1560  
QY 3794 AAATTCAGT 3803  
DB 1561 AAATTCAGT 1570

Search completed: March 5, 2004, 05:11:49  
Job time : 9628.23 secs

GenCore version 5.1.6	
Copyright (c) 1993 - 2004 CompuGen Ltd.	
OM nucleic - nucleic search, using sw model	
Run on:	March 4, 2004, 16:59:34 ; Search time 6050.99 Seconds
(without alignments)	
18847.119 Million cell updates/sec	
Title:	
Perfect score: 3819	
Sequence: 1 atggtgcccagcgccggc.....aggtgcaaacatgaacatga 3819	
Scoring table:	
IDENTITY_NUC	
Gapop 10.0 , Gapext 1.0	
Searched:	
27513289 seqs, 14931090276 residues	
Total number of hits satisfying chosen parameters:	
55026578	
Minimum DB seq length: 0	
Maximum DB seq length: 2000000000	
Post-processing: Minimum Match 0%	
Maximum Match 100%	
Listing first 45 summaries	
Database :	
EST:	
1: em_estba:**	
2: em_esthum:**	
3: em_estin:**	
4: em_estmu:**	
5: em_estov:**	
6: em_estpl:**	
7: em_estro:**	
8: em_estc:**	
9: gb_est1:**	
10: gb_est2:**	
11: gb_est3:**	
12: gb_est4:**	
13: gb_est5:**	
14: gb_est6:**	
15: em_estom:**	
16: em_estom:**	
17: em_estom:**	
18: em_estinv:**	
19: em_estpl:**	
20: em_estvrt:**	
21: em_estfun:**	
22: em_estmam:**	
23: em_estmus:**	
24: em_estpro:**	
25: em_estrod:**	
26: em_estphg:**	
27: em_estvrl:**	
28: gb_est1:**	
29: gb_est2:**	

13: gb_est4.*		14: gb_est5.*		15: em_estfun.*		16: em_estom.*		17: em_gss_hum.*		18: em_gss_inv.*		19: em_gss_pln.*		20: em_gss_vrt.*		21: em_gss_fun.*		22: em_gss_mam.*		23: em_gss_mus.*		24: em_gss_pro.*		25: em_gss_rod.*		26: em_gss_phg.*		27: em_gss_vrl.*		28: gb_gss1.*		29: gb_gss2.*	
Pred. No. is the number of results predicted by chance to have a		score greater than or equal to the score of the result being printed,		and is derived by analysis of the total score distribution.																													
SUMMARIES																																	
Result		Query		Score		Match		Length		DB		ID		Description																			
1		2940.6		77.0		3976		29		AY418587		Homo sapi																					
2		2507		65.6		3671		29		AY418588		Pan trogl																					
3		2507		65.6		3976		29		AY418589		Mus muscu																					
4		925.6		24.2		1201		13		BX440935		BX440935																					

ALIGNMENTS	
RESULT 1	3976 bp DNA linear GSS 17-DEC-2003
AY418587	
LOCUS	Homo sapiens ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION	AY418587
ACCESSION	AY418587
VERSION	AY418587.1 GI:39774547
KEYWORDS	GSS.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
	1 (bases 1 to 3976)
	Todd,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
	Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
	Adams,M.D. and Cargill,M.
	Inferring nonneutral evolution from human-chimp-mouse orthologous
	gene trios
	Science 302 (5652), 1960-1963 (2003)
JOURNAL	14671302
PUBMED	2 (bases 1 to 3976)
REFERENCE	Todd,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
AUTHORS	Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
	Adams,M.D. and Cargill,M.
	Direct Submission
TITLE	



JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT These sequences were made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
source Location/Qualifiers  
1..3976  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>3976  
/gene="ATRN"  
/locus\_tag="HCM6607"  
ORIGIN  
Query Match 77.0%; Score 2940.6; DB 29; Length 3976;  
Best Local Similarity 84.4%; Pred.No.0;  
Matches 2943; Conservative 0; Mismatches 546; Indels 0; Gaps 0;  
QY 315 TGTCAACGGCGTGTGCAACCTTGGACCGCGCAGTGCCTGCTCCCGCGCGGTGGGT 374  
Db 1 TGTCAACGGCGTGTGCAACCTTGGACCGCGCAGTGCCTGCTCCCGCGCGGTGGGT 60  
QY 375 GGGCGAGCAATCCAGCACTCGGGCGCGCTTACAGCTAAGTCTTCTGGGTTGT 434  
Db 61 GGGCGAGCAATCCAGCACTCGGGCGCGCTTACAGNNNNNNNNNNNNNNNNNN 120  
QY 435 GACAGATGACCTGGAAATATATAAACAACGAAGTGCAGCTGGCTCAITGAAGGACA 494  
Db 121 NNN 180  
QY 495 GCCAATAGAAATAGAGACTTCGTTTCAATCTTTGCTACAGAGTGTAGTGGACCA 554  
Db 181 GCCAATAGAAATAGAGACTTCGTTTCAATCAITTTGCTACAGAGTGTAGTGGACCA 240  
QY 555 TTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTGCAITTTAGTGGCT 614  
Db 241 TTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTGCAITTTAGNNNNN 300  
QY 615 CATTGCTCCTGAGAGATGGCAATGAGACTGTCCTGAGGTTGTGCCACATCAGTTA 674  
Db 301 NNN 360  
QY 675 TGCCTTGTGCTATTTTATGATGCTGCTTAAATTTTGACTGGAATTAATATCTTA 734  
Db 361 NNN 420  
QY 735 CAGTTTTCATATGTTCCAAATAACTGCTCAGGCGGAGGAGTGTAAAGATCAGTAATAG 794  
Db 421 NNNTTTTCATATGTTCCAAATAACTGCTCAGGCGGAGGAGTGTAAAGATCAGTAATAG 480  
QY 795 CAGCGAACTGTTGAATGTAATGTTCTGAAACTGGAAGAGTGAAGATGTGACATCC 854  
Db 481 CAGCGAATCTGTTGAATGTAATGTTCTGAAACTGGAAGAGTGAAGATGTGACATCC 540  
QY 855 TCACGTACAGACAACTGTGGTTTCTCTCAGGAGGATCTGCAATTCAGTGTATGTCAG 914  
Db 541 TCACGTACAGACAACTGTGGTTTCTCTCAGGAGGATCTGCAATTCAGTGTATGTCAG 600  
QY 915 AGGATGCTCCTGTTCTCAGACTGGCGGGTCTGGATGTTTCAGTTCCTGTACAGCTAA 974  
Db 601 AGGATGCTCCTGTTCTCAGACTGGCGGGTCTGGATGTTTCAGTTCCTGTACAGCTAA 660  
QY 975 CCAGTCATTTTGGACTCGAGAGGAATATCTTAAGCTCCCGAGAGCACTCATAA 1034  
Db 661 CCAGTCATTTTGGACTCGAGAGGAATATCTTAAGCTCCCGAGAGCACTCATAA 720  
QY 1035 AGCTGTGCTCAATGAAACATTAATGTGGTGTGAGGATATATGTTCAACCACTCAGA 1094  
Db 721 AGCTGTGCTCAATGAAACATTAATGTGGTGTGAGGATATATGTTCAACCACTCAGA 780  
QY 1095 TTATAACATGTTCTAGGATGACCTTGTCTTACGGAGTGGCTTCCACTAACCGTTT 1154  
Db 781 TTATAACATGTTCTAGGATGACCTTGTCTTACGGAGTGGCTTCCACTAACCGTTT 840

QY 1155 TGTGAACAATGTGGTTGTTAGATATGTGTCATCTTTGGCAATATATACAAGGATAAAATTTA 1214  
Db 841 TGTGAACAATGTGGTTGTTAGATATGTGTCATCTTTGGCAATATATACAAGGATAAAATTTA 900  
QY 1215 CATGTATGGAGGAAATATGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTC 1274  
Db 901 NNN 960  
QY 1275 CATTCAATAATGAGTCAATGGGTGTGTGACCCCTAAGGCAAGAGCAGATATGCACTGT 1334  
Db 961 NNN 1020  
QY 1335 TGGGCACTCTGCACACATCTTACATGAAGATGCCGAGTGTGTCATGCTGCTCATCTT 1394  
Db 1021 NNN 1080  
QY 1395 TGTCTCATCTGCCCTCTCTATGGATATATAAGCAATGTGCGAGGAATATGATTTGGATAAGAA 1454  
Db 1081 NNN 1140  
QY 1455 CACATGGAGTATATTACACACCCAGGTCCTTGTGCAAGGGGTTACGGCCATAGCAG 1514  
Db 1141 CACATGGAGTATATTACACACCCAGGTCCTTGTGCAAGGGGTTACGGCCATAGCAG 1200  
QY 1515 TGTTTACGACCATAGGACAGGGCCCTATACGTTTCTGTTGGCTACAAGGCTTTCAGTGC 1574  
Db 1201 TGTTTACGACCATAGGACAGGGCCCTATACGTTTCTGTTGGCTACAAGGCTTTCAGTGC 1260  
QY 1575 CAATAAGTACCGCTTGCAGATGATCTCTACGATATGATGTGGATACCCAGATGTGGAC 1634  
Db 1261 CAATAAGTACCGCTTGCAGATGATCTCTACGATATGATGTGGATACCCAGATGTGGAC 1320  
QY 1635 CATTCTTAAGGACAGCGGATTTTCCGTTTACTTGGACACAGCTGTGATAGTACAGTGGAC 1694  
Db 1321 CATTCTTAAGGACAGCGGATTTTCCGTTTACTTGGACACAGCTGTGATAGTACAGTGGAC 1380  
QY 1695 CATGCTGTTTGGGGGAAACACACAAATGACATCTATGAGCCATGGGCCCAATG 1754  
Db 1381 CATGCTGTTTGGGGGAAACACACAAATGACATCTATGAGCCATGGGCCCAATG 1440  
QY 1755 CTTCTCTCAGATTTTCATGCGCTATGACATTCCTGTGACCGCTGTGCTGCTCCAG 1814  
Db 1441 CTTCTCTCAGATTTTCATGCGCTATGACATTCCTGTGACCGCTGTGCTGCTCCAG 1500  
QY 1815 ACCTGATCTCCACCATGATGTCAACAGATTTTGGCCATTCAGCAGTCTTACACACAGCAC 1874  
Db 1501 NNN 1560  
QY 1875 CATGATGTTTGGTGGTTTCAATAGTCTCTCTCAGCGACATCTCTGGTATTCACCTC 1934  
Db 1561 CATGATGTTTGGTGGTTTCAATAGTCTCTCTCAGCGACATCTCTGGTATTCACCTC 1620  
QY 1935 GGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTAGCAGCAGACCTGGTATTCG 1994  
Db 1621 GGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTAGCAGCAGACCTGGTATTCG 1680  
QY 1995 GTGTGTGGAACACACAGGTCCTCTCAGTGTATCTCTGGGCGCTGGCAACTGTATGAACA 2054  
Db 1681 GTGTGTGGAACACACAGGTCCTCTCAGTGTATCTCTGGGCGCTGGCAACTGTATGAACA 1740  
QY 2055 AGAAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCATGACAGATGTGA 2114  
Db 1741 AGAAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCATGACAGATGTGA 1800  
QY 2115 CCAGCACACAGATTTGTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTGAATGA 2174  
Db 1801 CCAGCACACAGATTTGTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTGAATGA 1860  
QY 2175 CCATTGTGTCCTCCAGGACCACTGCTCAGAGGCGCAGATCTCCATTTTATAGTATGA 2234  
Db 1861 CCATTGTGTCCTCCAGGACCACTGCTCAGAGGCGCAGATCTCCATTTTATAGTATGA 1920



2235 GAATGCCCCAAGGATTAACCCCATGTAATACTGTAACAAGAGACGAGCTGCGAGGAGCTG 2294  
Db  
1921 GAATGCCCCAAGGATTAACCCCATGTAATACTGTAACAAGAGACGAGCTGCGAGGAGCTG 1980  
Qy  
2295 TGCCTCGGACAGAACTGCGAGTGGGAGCCCGGATCAGAGTGCATTCGCCCTGCCCCG 2354  
Db  
1981 TGCCTCGGACAGAACTGCGAGTGGGAGCCCGGATCAGAGTGCATTCGCCCTGCCCCG 2040  
Qy  
2355 AAATATCTGTGGCAATGCGCTGGCAATTTGGTTGGAACTCATGTTGAAAATTAATACTGC 2414  
Db  
2041 AAATATCTGTGGCAATGCGCTGGCAATTTGGTTGGAACTCATGTTGAAAATTAATACTGC 2100  
Qy  
2415 CAAGGAGATTTACAGAACTGCTAAATTTGTTCTGTAGGAACCAACATGCCCTTTGGCTTC 2474  
Db  
2101 CAAGGAGATTTACAGAACTGCTAAATTTGTTCTGTAGGAACCAACATGCCCTTTGGCTTC 2160  
Qy  
2475 TCTTTACAAACCCAGAGAGTAGAATTTGCTTTAAGCAGCTGCGAATAATGCACTCATC 2534  
Db  
2161 TCTTTACAAACCCAGAGAGTAGAATTTGCTTTAAGCAGCTGCGAATAATGCACTCATC 2220  
Qy  
2535 TCAGAGCATGTCGAAGCTCAACCTTAACCCCATGCGCTCGGCTTCGGAAGATCAATGTGTC 2594  
Db  
2221 TCAGAGCATGTCGAAGCTCAACCTTAACCCCATGCGCTCGGCTTCGGAAGATCAATGTGTC 2280  
Qy  
2595 CTACTGTTGCTGGCAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCCGTC 2654  
Db  
2281 CTACTGTTGCTGGCAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCCGTC 2340  
Qy  
2655 TGAGCCCAAGTGAATCTGGAATTTGGAATTTATCAGAACCCAGTACTCGGGAGCTGAA 2714  
Db  
2341 TGAGCCCAAGTGAATCTGGAATTTGGAATTTATCAGAACCCAGTACTCGGGAGCTGAA 2400  
Qy  
2715 GGCTGCAACCTGCATCAACCACTCAATGTGAGTGTCTGCAAGAGCTGCAACACACAG 2774  
Db  
2401 GGCTGCAACCTGCATCAACCACTCAATGTGAGTGTCTGCAAGAGCTGCAACACACAG 2460  
Qy  
2775 TGCTAAGCAGTGCAGGACACCATGTCCTTGAGACAGCATGTGAGATTCGACACAGCG 2834  
Db  
2461 TGCTAAGCAGTGCAGGACACCATGTCCTTGAGACAGCATGTGAGATTCGACACAGCG 2520  
Qy  
2835 CAGCTCTGAGTGCATGTGTCAGCAACATGAACAGTGTGTCACCTGCAATGCTATGT 2894  
Db  
2521 CAGCTCTGAGTGCATGTGTCAGCAACATGAACAGTGTGTCACCTGCAATGCTATGT 2580  
Qy  
2895 GGCTCTCTCTCTCTTTGGCCAGTGTATGGAATGTATACAGATGAGCACCTGCCCTCTGA 2954  
Db  
2581 GGCTCTCTCTCTCTTTGGCCAGTGTATGGAATGTATACAGATGAGCACCTGCCCTCTGA 2640  
Qy  
2955 AAATGTTTCAAGCTACTGTACTGTAGTCAATGCTTGAGAACCAACAGCTGTGCTGGTG 3014  
Db  
2641 AAATGTTTCAAGCTACTGTACTGTAGTCAATGCTTGAGAACCAACAGCTGTGCTGGTG 2700  
Qy  
3015 TACTGATCCCAAGCAATACCTGCAAGGGAATGATGATAGAGGTTTCTTAAAGAACAGT 3074  
Db  
2701 TACTGATCCCAAGCAATACCTGCAAGGGAATGATGATAGAGGTTTCTTAAAGAACAGT 2760  
Qy  
3075 GAAGATGCTTTGCGAAGCCCTTACAGGAAATTTCTATCCAGAGCCCTGCTCAATTCAG 3134  
Db  
2761 GAAGATGCTTTGCGAAGCCCTTACAGGAAATTTCTATCCAGAGCCCTGCTCAATTCAG 2820  
Qy  
3135 CATGTTCTAGAGACAGCATCACTGCTTTTCAATTCATTCAGTGCAGCTTCCCAATG 3194  
Db  
2821 CATGTTCTAGAGACAGCATCACTGCTTTTCAATTCATTCAGTGCAGCTTCCCAATG 2880  
Qy  
3195 CAACGGCCACAGTAATCAATCAAGCATCTGTGAGAAAGTGTGAGAACCTTGACAC 3254  
Db  
2881 CAACGGCCACAGTAATCAATCAAGCATCTGTGAGAAAGTGTGAGAACCTTGACAC 2940  
Qy  
3255 AGGCAAGCACTGCGAGACCTGCATATCTGGTTTCTACGGTATGCCCAATGAGGGAA 3314  
Db  
2941 AGGCAAGCACTGCGAGACCTGCATATCTGGTTTCTACGGTATGCCCAATGAGGGAA 3000  
Qy  
3315 ATGTCAGCCATGCAAGTGAATGGGACGCGCTCTGTGCAACCAACACGCGCAAGTG 3374

## RESULT 2

AY418588

LOCUS

DEFINITION

Pan troglodytes ATRN gene, 3671 bp DNA linear GSS 17-DEC-2003

genomic survey sequence.

ACCESSION

AY418588

VERSION

AY418588.1

SOURCE

Pan troglodytes (chimpanzee)

ORGANISM

Pan troglodytes

REFERENCE

1 (bases 1 to 3671)

AUTHORS

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,

Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,

Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,

Adams, M.D. and Cargill, M.

TITLE

Inferring nonneutral evolution from human-chimp-mouse orthologous

gene trios

JOURNAL

Science 302 (5652), 1960-1963 (2003)

PUBMED

14671302

REFERENCE

2 (bases 1 to 3671)

AUTHORS

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,

Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,

Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,

Adams, M.D. and Cargill, M.

TITLE

Direct Submission

JOURNAL

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

Rockville, MD 20850, USA

COMMENT

These sequences were made by sequencing genomic exons and ordering

them based on alignment.

FEATURES

Location/Qualifiers

source

1..3671

/organism="Pan troglodytes"

/mol\_type="genomic DNA"

/db\_xref="taxon:9598"

&lt;1..&gt;3671

gene

ORIGIN		/gene="ATRN" /locus_tag="HCM6607"	
Query Match		65.6%;	Score 2507; DB 29; Length 3671;
Best Local Similarity		76.0%;	Pred. No. 0;
Matches 2516;		Conservative 0;	Mismatches 793; Indels 0; Gaps 0;
QY	495	GCCAAATGAATGAATGAGACTTCGTTTCAATCAATTTGCTACAGAGTGTAGTTGGACCA	554
DB	1	GCCAAATGAATGAATGAGACTTCGTTTCAATCAATTTGCTACAGAGTGTAGTTGGACCA	60
QY	555	TTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTCTGCAATTTAGTGGCCT	614
DB	61	TTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTCTGCAATTTAGTGGCCT	120
QY	615	CATTGTTCTGACAGAGATGGAATGAGACTGTCCCTGAGGTGTGTGCCACATCAGGTGA	674
DB	121	NN	180
QY	675	TGCTTGTGCTGCAATTTTATGATGCTGCTGTATAATTTGATGGAATTAATTAATTA	734
DB	181	NN	240
QY	735	CAGTTTTCATATGCTCCAAATAACTGCTCAGGCGGAGGAGTGAAGATCAGTAATAG	794
DB	241	NN	300
QY	795	CAGCGAAACTGTGTAATGTAATGTTCTGAAACTGGAAGAGTGAAGCATGTGCAATTC	854
DB	301	CAGCGAATCTGTGAATGTAATGTTCTGAAACTGGAAGAGTGAAGCATGTGCAATTC	360
QY	855	TCAGTGTACAGACACTGTGTTTCTCATCGAGGATCTGCAATTCAGTGTAGTGTGAG	914
DB	361	TCAGTGTACAGACACTGTGTTTCTCATCGAGGATCTGCAATTCAGTGTAGTGTGAG	420
QY	915	AGGATGCTCCTGCTTCTCAGACTGGCAGGGTCTGGATGTTTCAGTCTCTGTACCACTAA	974
DB	421	AGGATGCTCCTGCTTCTCAGACTGGCAGGNNCTGGATGTTTCAGTCTCTGTACCACTAA	480
QY	975	CGAGTCATTTGGACTCGAGAGGAATATTCTAACTTAAAGTCCCGAGAGCATCTATAA	1034
DB	481	CGAGTCATTTGGACTCGAGAGGAATATTCTAACTTAAAGTCCCGAGAGCATCTATAA	540
QY	1035	AGCTGTGTCATGGAACAATTATGTGGTGTGTTGGAGGATATATGTTCAACCACTCAGA	1094
DB	541	AGCTGTGTCATGGAACAATTATGTGGTGTGTTGGAGGATATATGTTCAACCACTCAGA	600
QY	1095	TTATAACATGTTTACGGTATGACCTTGTCTTACGGAGTGGCTTCCACTAAACCGTTC	1154
DB	601	TTATAACATGTTTACGGTATGACCTTGTCTTACGGAGTGGCTTCCACTAAACCGTTC	660
QY	1155	TGTGAACAATGTTGTTAGATATGTTGCTATCTTCTGCAATTATACAGGATATAATTTA	1214
DB	661	TGTGAACAATGTTGTTAGATATGTTGCTATCTTCTGCAATTATACAGGATATAATTTA	720
QY	1215	CATGTATGGAGGAAAAATTGATCCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTC	1274
DB	721	NN	780
QY	1275	CATTCAATGAGTCATGGGTGTTGTTGACCCCTAAGCCAAAGAGCAGATGACAGTGGT	1334
DB	781	NN	840
QY	1335	TGGGCACTCTGCACACATTTGTAACACTGAAGAAATGGCCGAGTGTCTATGCTGTATCTT	1394
DB	841	NN	900
QY	1395	TGGTCACTGCCCTCTATGATGATATAAGCAATGTGAGGAATATGATTTGGATAGAA	1454
DB	901	NN	960
QY	1455	CACATGGAGTATATACACACCCAGGGTGCCCTTGTGCAAGGGGTTTACGGCCATAGCAG	1514
DB	961	CACGTGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGTTTACGGCCATAGCAG	1020
QY	1515	TGTTTACGACCATAGGACCCAGGGCCCTATACGTTTATGTTGCTTCAAGGCTTTTCAAGTGC	1574
DB	1021	TGTTTACGACCATAGGACCCAGGGCCCTATACGTTTATGTTGCTTCAAGGCTTTTCAAGTGC	1080
QY	1575	CAATAAGTACCGGCTTTCAGATGATCTCTACCGATATGATGTGGATATCCAGATGTGGAC	1634
DB	1081	CAATAAGTACCGGCTTTCAGATGATCTCTACCGATATGATGTGGATATCCAGATGTGGAC	1140
QY	1635	CATTCTTAAAGGACAGCCGATTTTCGTTTACTTGGACACAGCTGTGATAGTGGGAC	1694
DB	1141	CATTCTTAAAGGACAGCCGATTTTCGTTTACTTGGACACAGCTGTGATAGTGGGAC	1200
QY	1695	CATGCTGTGTTTGGGGGAAACACACAAATGACACATCTATGAGCCATGGCGCCAAATG	1754
DB	1201	CATGCTGTGTTTGGGGGAAACACACAAATGACACATCTATGAGCCATGGCGCCAAATG	1260
QY	1755	CITCTCTTCCAGATTTTATGCGCTTATGACATTTGCTGTGACCGCTGTCAGTCTTCCAG	1814
DB	1261	CITCTCTTCCAGATTTTATGCGCTTATGACATTTGCTGTGACCGCTTCCAGATTTTCCAG	1320
QY	1815	ACCTGATCTCCACCATGATGTCAAACAGATTGCGCATTCAGCAGTCTTACACACAGCAC	1874
DB	1321	NN	1380
QY	1875	CATGATATGTTTGGTGGTTCCTCAATGATCTCTCTCAGCGACATCTCTGTTATTCACCTC	1934
DB	1381	CATGATATGTTTGGTGGTTCCTCAATGATCTCTCTCAGCGACATCTCTGTTATTCACCTC	1440
QY	1935	GGAACTGTGATGCGCATCGAGTGAAGCGCTTGTGTAGCAGCAGACCTGGTATTCG	1994
DB	1441	GGAACTGTGATGCGCATCGAGTGAAGCGCTTGTGTAGCAGCAGACCTGGTATTCG	1500
QY	1995	GTGTGTGTGAAACACAGGTCCTCTCAGTGTATCTCTGTTGGCGCTGGCAACTGATGAACA	2054
DB	1501	GTGTGTGTGAAACACAGGTCCTCTCAGTGTATCTCTGTTGGCGCTGGCAACTGATGAACA	1560
QY	2055	AGAAGAAAGTTAAATCAGATGTTTTCCTCAAGAACTCTTGACCATGACAGATGGA	2114
DB	1561	AGAAGAAAGTTAAATCAGATGTTTTCCTCAAGAACTCTTGACCATGACAGATGGA	1620
QY	2115	CCAGCACACAGATTTGTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTGAATGA	2174
DB	1621	CCAGCACACAGATTTGTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTGAATGA	1680
QY	2175	CCATTGTGTCCTCAGGAAACACAGCTGTCTCAGAGCCAGATCTCCATTTTGTAGTATGA	2234
DB	1681	CCATTGTGTCCTCAGGAAACACAGCTGTCTCAGAGCCAGATCTCCATTTTGTAGTATGA	1740
QY	2235	GAATTGCCCCCAAGGATAACCCCATCTACTGTAAACCAAGAACCCAGCTGCAGAGCTG	2294
DB	1741	GAATTGCCCCCAAGGATAACCCCATCTACTGTAAACCAAGAACCCAGCTGCAGAGCTG	1800
QY	2295	TGCCCTGGAACCAAGAACTGCCAGTGGAGCCCGGAATCAGGAGTGCATTTGCCCGCA	2354
DB	1801	TGCNNGGACCAAGAACTGCCAGTGGAGCCCGGAATCAGGAGTGCATTTGCCCGCA	1860
QY	2355	AAATATCTGTGTCATTTGCTGTCATTTGTTGGAACTCATGTTTGAATAATCTACTGTC	2414
DB	1861	AAATATCTGTGTCATTTGCTGTCATTTGTTGGAACTCATGTTTGAATAATCTACTGTC	1920
QY	2415	CAAGGAGAAATTTATGACAAATGCTTAAATTTCTGTAGGAACCAACATGCCCTTTTGGCTTC	2474
DB	1921	CAAGGAGAAATTTATGACAAATGCTTAAATTTCTGTAGGAACCAACATGCCCTTTTGGCTTC	1980
QY	2475	TCCTTCAACCCAGAGAGAGTGAATTTGCTTTAAGCAGCTGCGAATAATGTCAGTCAATC	2534
DB	1981	TCCTTCAACCCAGAGAGAGTGAATTTGCTTTAAGCAGCTGCGAATAATGTCAGTCAATC	2040
QY	2535	TCAGAGCATGTCCAGCTCACCTTAAACCCCATGGCTCGGCCTTCGGAGATCAATGTGTC	2594

2041 TCNNNNNNNTCCAACTCACCTTAACCCCATGGGTGGGCTCGGAGATCAATGTGTC 2100  
2595 CTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCCGTC 2654  
2101 CTACTGCTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCCGTC 2160  
2655 TGAGCCCACTGATGCTGGATCTGTGGAATTTATCAGAACCCAGTACTCGGGGACTGAA 2714  
2161 TGAGCCCACTGATGCTGGATCTGTGGAATTTATCAGAACCCAGTACTCGGGGACTGAA 2220  
2715 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACCCAG 2774  
2221 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACCCAG 2280  
2775 TGCTAAGCAGTCCCGGACACCATGTGCTTGGAGACAGCATGTGGAGATTCACCCAGCGG 2834  
2281 TGCTAAGCAGTCCCGGACACCATGTGCTTGGAGACAGCATGTGGAGATTCACCCAGCGG 2340  
2835 CAGCTCTGAGTGCATGTGGTGCAGCAACATGAGCAGTGTGGACTCCAATGCTATGT 2894  
2341 CAGCTCTGAGTGCATGTGGTGCAGCAACATGAGCAGTGTGGACTCCAATGCTATGT 2400  
2895 GGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGTCCGCCCTGA 2954  
2401 GGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGTCCGCCCTGA 2460  
2955 AAATGTTCAAGCTACTGACTGCTGATGCTGAGTGTGGAGCAACAGGCTGTGGCTGTG 3014  
2461 NNN 2520  
3015 TACTGATCCAGCAATACTGGCAAGGAAATGCATAGAGGTTCTCTATAAGGACCAGT 3074  
2521 NNN 2580  
3075 GAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATCCAG 3134  
2581 NNN 2640  
3135 CATGCTGTAGAGGACAGAGATACAACTGGTCTTCTTCTACTGTCAGCTGTCGAATG 3194  
2641 NNN 2700  
3195 CAACGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTCAGCAC 3254  
2701 CAACGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTCAGCAC 2760  
3255 AGGCAAGCATCGAGACCTGATCTGGCTTCTACGGTGTATCCCAATGGAGGAA 3314  
2761 AGGCAAGCATCGAGACCTGATCTGGCTTCTACGGTGTATCCCAATGGAGGAA 2820  
3315 ATGTCAGCCATCAAGTCAATGGGACGCGTCTGTGCAACACCAACAGCGGCAAGTG 3374  
2821 NNN 2880  
3375 CTTCTGACCAACCAAGGGCGTCAAGGGGACGAGTCCAGCTATGTGAGGTAGAAAATCG 3434  
2881 CTTCTGACCAACCAAGGGCGTCAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2940  
3435 ATACCAAGAAACCCCTCTCAGAGGAACATGTTATATCTCTCTTATGACTATCAGTT 3494  
2941 ATACNN 3000  
3495 CACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATCAATTTTGTGGCTAC 3554  
3001 CACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATTAATTTTGTGGCTAC 3060  
3555 TCTTGACGAACAAAACAGGATTTGGACATGTTTCAATGCTTCAAGAAATTTCAACT 3614  
3061 TCTTGACGAACAAAACAGGATTTGGACATGTTTCAATGCTTCAAGAAATTTCAACT 3120  
3615 CACATCACTGGCTGGCTGTCAGTCTGCAACCCAGGCTGGAGAGAGATGCTGT 3674  
3121 CACATCACTGGCTGGCTGTCAGTCTGCAACCCAGGCTGGAGAGAGATGCTGT 3180

3675 TGTTCAAAACCAACCATTAAGAGTACAAAGATGTTTCTCTAAATGAGAAGTTGATTT 3734  
3181 TGTTCAAAACCAACCATTAAGAGTACAAAGATGTTTCTCTAAATGAGAAGTTGATTT 3240  
3735 TCGCAACCCACCAAAATATCACTTTCTTTTATGTAGTAAATTTTCCCTGGCCCATCAA 3794  
3241 TCGCAACCCACCAAAATATCACTTTCTTTTATGTAGTAAATTTTCCCTGGCCCATCAA 3300  
3795 AATTCAAGT 3803  
3301 AATTCAAGT 3309

RESULT 3  
AY418589  
LOCUS  
DEFINITION  
Mus musculus ATRN gene, 3976 bp DNA linear GSS 17-DEC-2003  
genomic survey sequence.

ACCESSION  
AY418589  
VERSION  
AY418589.1 GI:39774549  
KEYWORDS  
GSS.  
SOURCE  
Mus musculus (house mouse)  
ORGANISM  
Mus musculus

REFERENCE  
1 (bases 1 to 3976)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.

TITLE  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios

JOURNAL  
Science 302 (5652), 1960-1963 (2003)  
PUBMED  
14671302  
REFERENCE  
2 (bases 1 to 3976)  
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,  
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,  
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,  
Adams,M.D. and Cargill,M.

TITLE  
Direct Submission:  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA

COMMENT  
These sequences were made by sequencing genomic exons and ordering  
them based on alignment.

FEATURES  
Location/Qualifiers  
source  
1..3976  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
<1..3976  
/gene="ATRN"  
/locus\_tag="HCM6607"

ORIGIN

Query Match 65.6%; Score 2507; DB 29; Length 3976;  
Best Local Similarity 76.6%; Pred. No. 0;  
Matches 2672; Conservative 0; Mismatches 817; Indels 0; Gaps 0;

315 TGTCAACGGCGTCCGTGCAACCTGCGACCGCCAGTGGCTGTGCCCGCGCTGGT 374  
1 TGTCAACGGCGCGCTGCAACCTGCGACCGCCAGTGGCTGTGCCCGCGCTGGT 60  
375 GGGCGAGCAATGCCAGCACTGCGGGGCGCTTCAGACTAACTGGATCTTCTGGGTTGT 434  
61 GGGCGAGCAATGCCAGCACTGCGGGGCGCTTCAGNNNNNNNNNNNNNNNNNNNN 120  
435 GACAGATGGAGCTGGAATTAATAACAAACGAAGTGCAGCTGGCTCATTTGAAGACA 494  
121 NNN 180  
495 GCCAAATAGATAATAGAGACTTCGTTTCAATCATTTGCTACAGAGTGTAGTGGACCA 554

Db 181 GCCAAATAGAAATATGAGACTTCGGCTTCAACCATTTTGTCTACAGATGTAGCTGGGACCA 240  
Qy 555 TTATATATGTTATGATGGGAGCTCAATTTATGACCGCTAGTGTGCTGATTTAGTGGCT 614  
Db 241 TTATATATGTTATGATGGGAGCTCAATCTACGCACTCTGATGCTGTGCTTAGNNNNN 300  
Qy 615 CATTTGCTCTGAGAGATGGCAATGAGACTGTCCTGAGGTGTGCCACATCAGTTA 674  
Db 301 NNN 360  
Qy 675 TGCCTTGTGTCATTTTTTTTAGTGATGCTCTTATAATTTGACTGCAATTAATTAATTA 734  
Db 361 NNN 420  
Qy 735 CAGTTTGTATGTGTCGAAATACTGCTCAGCGCGAGAGAGTGTAAAGTACAGTAATAG 794  
Db 421 NNNTTTGTACATGTGTCGGAATAATGTCTCAGCGCGAGAGAGTGTAAAGCAGTAA 480  
Qy 795 CAGCGAAACTGTTGAATGTGAATGTTCTGAAAACTGGAAGGTGAAGCATGTGCAATTC 854  
Db 481 CAGCAGCGCTGTGAGTGTGAATGTTCTGAAAACTGGAAGGTGAAGCATGTGCAATTC 540  
Qy 855 TCACTGTACAGCAAACTGTGTTTCTCATCGAGGCACTGCAATCAAGTGTGTCAG 914  
Db 541 TCACTGTACAGCAAACTGTGTTTCTCATCGAGGCACTGCAATCAAGTGTGTCAG 600  
Qy 915 AGGATGCTCTGCTCTCAGATGCGCAGGCTCTGATGTTCAAGTTCCTGTACCACTAA 974  
Db 601 AGGATGCTCTGCTCTCAGATGCGCAGGCTCTGATGTTCAAGTTCCTGTACCACTAA 660  
Qy 975 CCAGTCATTTTGGACTCGAGAGGAATATTTCTAACTTAAAGTCTCCAGAGCATCTCA 1034  
Db 661 CCAGTCATTTTGGACTCGAGAGGAATATTTCTGATTTAAAGCTTCCAGAGCTCTCA 720  
Qy 1035 AGCTGTGTCATGGAACATATGATGTTGTTGGAGATATATGTTCAACCACTCAGA 1094  
Db 721 AGCTGTGTCATGGAACATATGATGTTGTTGGAGATATATGTTCAACCACTCAGA 780  
Qy 1095 TTATAACATGTTCTAGCGTATGACTTCTTAGGAGTGGCTTCCACTTAAACCGTTC 1154  
Db 781 TTACAGCATGTTTATGCGTATGACTGACTTCTAGGGAATGGCTTCCACTTAAACCTC 840  
Qy 1155 TGTGAACATGTTGTTGATATGTTGTCATTTTGGCATATATACAGGATATAAATTTA 1214  
Db 841 TGTGAACATGTTGTTGATATGTTGTCATTTTGGCATATATACAGGATATAAATTTA 900  
Qy 1215 CATGTATGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAGTGTGAGAGTTTCA 1274  
Db 901 NNN 960  
Qy 1275 CATTCATAATGAGTCATGGGTGTTGTCACCCCTAAGCAAGAGCAGTATGCAAGTGT 1334  
Db 961 NNN 1020  
Qy 1335 TGGGCACCTGACACATGTTTACACTGAAGATGGCCGAGTGTGCTGCTGCTCATCTT 1394  
Db 1021 NNN 1080  
Qy 1395 TGGTCACCTGCCCTCTATGATATATAAGCAATGTGCAAGATATGATTTGGATTAAGAA 1454  
Db 1081 NNN 1140  
Qy 1455 CACATGGAGTATATTTACACACCGGGTGCCTGTGCAAGGGGTATCGGCCATAGCAG 1514  
Db 1141 CACATGGAGTATATTTACATCTCAGGGTGTCTTGTGCAAGGGGTATGCGCCACAGTAG 1200  
Qy 1515 TGTTTACACCAATAGGACACAGGGCCCTATAGCTTCACTGGTGTGCTACAGGGCTTTCAGTGC 1574  
Db 1201 TGTTTATGATGACAGACCAAGGCTCTGTAAGTTCACTGGTGTGCTACAGGGCTTTCAGTGC 1260  
Qy 1575 CAATAAGTACCGGCTTGCAGATGATCTCTACCGATGATGTGATACCCAGATGTGAC 1634  
Db 1261 CAACAAATACCGGCTTGCAGATGATGCTCTACAGATACAGATGTGATCTCAGATGTGAC 1320

Qy 1635 CATTTTAAAGACAGCGGATTTTTTCCGTACTTGTGCAACAGCTGTGATAGTGTGGAAC 1694  
Db 1321 CATTTTAAAGACAGCGGATTTTTTCCGTACTTGTGCAACAGCTGTGATAGTGTGGAAC 1380  
Qy 1695 CATGCTGCTGTTTGGGGAAACACACATGATGACATCTATGAGCATGGGCCAAATG 1754  
Db 1381 CATGCTGCTGTTTGGAGGAAACACACATGATGACATCTTCCATGAGCCAGGTCGCAATG 1440  
Qy 1755 CTTCTCTTCAGATTTTATGCGCTTATGACATGCTGTGACCGCTGTGCTGCTTCCAG 1814  
Db 1441 CTTCTCTTCAGATTTTATGCGCTTATGACATGCTGNNNNNNNNNNNNNNNNNNNNNNNN 1500  
Qy 1815 ACCTGATCTCCACATGATGTCAACAGATTTGGCAATTCAGCAGTCTTACACACAGCAC 1874  
Db 1501 NNN 1560  
Qy 1875 CATGTATGTTTCCGTGTTTCAATAGTCTCTCTCAGCAGCATCTCTGTATTCACCTC 1934  
Db 1561 CATGTATGTTTCCGTGTTTCAACAGCTCTCTCTCAGTGTGCTTGTCTTTACCTC 1620  
Qy 1935 GGAACAGTGTGATGCGCATGCGAGTGAAGCGCTTGTAGCAGCAGGACCTGTGTTTCG 1994  
Db 1621 GGAGCAGTGTGATGSCACACCGCAGTGAAGCTGTTGTGGCAGCAGGACCTGGTATCG 1680  
Qy 1995 GTGTGTGTGAAACACAGGCTGTCTCAGTGTATCTCTGTGGCGCTGGCAACTGATGAAC 2054  
Db 1681 GTGTGTGTGAAACACAGTGTCTCAGTGTATCTCTGTGGAGTGGCAACTGATGAAC 1740  
Qy 2055 AGAAGAAAGTATAAATCAGAAATTTTTTCCAAAGAACTCTTGACCATGACAGATGTA 2114  
Db 1741 AGCAGAAAGTATAAATCAGAGTGTTTTTTCTAAAGAACTTGACCATGACAGATGTA 1800  
Qy 2115 CCAGCACACAGATGTGATGCTGTACAGCAGCAACCAATGACTGCTGCTGTGCAATGA 2174  
Db 1801 CCAGCACACAGATGTGATGCTGTACAGCAGCAACCAATGACTGCTGCTGTGCAATGA 1860  
Qy 2175 CCATTTGTCTCCAGGAAACACAGCTGTCTCAGAAAGCCAGATCTCCATTTTGTAGTATGA 2234  
Db 1861 TCACTGTGCTGTGAAACACAGCTGTGACAGAGCCAGATCTCCATTTGCAAGTATGA 1920  
Qy 2235 GAATTCGCCCAAGGATTAACCCATGTACTGTATACAGAGACCAAGCTGTGAGAGCTG 2294  
Db 1921 GAGTTGCCCAAGGATTAACCCATGTACTGTATGCAATTAAGAAACCAAGCTGTGAGAGCTG 1980  
Qy 2295 TGCCCTGGACCAAGCTGCGAGTGGGAGCCCGGAATCAGAGTGTGATGCTGCTGCCGA 2354  
Db 1981 TGCCCTGAGCAGAACTGCGAGTGGGAGCCCGGAATCAGAGTGTGATGCTGCTGCCGA 2040  
Qy 2355 AAATATCTGTGGCATTTGGTGTGCAATTTGTTGAAACTCATGTTTGAATAATTAAGTGTG 2414  
Db 2041 AAATATCTGTGGCAATTTGGTGTGCAATTTGTTGAAACTCATGTTTGAATAATTAAGTGTG 2100  
Qy 2415 CAAGGAGATTTATGACATGCTTAAATGTTCTGTAGGACCAACATGCCCTTTGGCTTC 2474  
Db 2101 TAAGGAGATTTATGACATGCTTAAATGTTCTGTAGGACCAACATGCCCTTTGGCTTC 2160  
Qy 2475 TCTTACACCCGAGAGAGTGAATTTGTTCTTAAAGCAGCTGCGAATAATCAGTGTATC 2534  
Db 2161 CCTCACATCCGAGAGAGTGGAGTTTGTCTTAAAGCAGCTTGCATTAATCAATCATC 2220  
Qy 2535 TCAGACATGTCAGAGCTCACCTTAACCCCATGGGTGGCTTCGAGAGTCAATGTGTC 2594  
Db 2221 TCAAGATGTCAGAGCTCACCTTAACCCCATGGGTGGCTTCGAGAGTCAATGTGTC 2280  
Qy 2595 CTACTGCTCTGGAGATATGTCCTCCCAATTTTACAAATAGTTTACTACAGTGTGATCCCTC 2654  
Db 2281 TTACTGCTCTGGAGATATGTCCTCAATTTCAAAATAGTTTGTGCTGAGTGTGATGCCATC 2340  
Qy 2655 TGAGCCAGTGTGATGCTGGAATTTCTGGAATTTTATCAGAACCCAGTACTCGGGAGTGA 2714  
Db 2341 TGAGCCAGTGTGATGCTGGAATTTCTGGAATTTTATCAGAACCCAGTACTCGGGAGTGA 2400

QY	2715	GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGGCTGCAAAACACAG	2774
Db	2401	GGCTGCAACCTGCATCAACCCCTCTCAATGGCAGCGTCTGTGAAGGGCTGCAAAACACAG	2460
QY	2775	TGCTAAGCAGTGCAGGACACCATGTGCTTGCAGACAGCATGTGGAGATTGCACACGCG	2834
Db	2461	TGCCAAGCAGTGCAGGACACCATGTGCTTGCAGACAGCATGTGGAGATTGCACACGCG	2520
QY	2835	CAGCTCTGAGTGCATGTGGTGCAGCAACATGAACAGTGTGTGACCTCCATGCTTATGT	2894
Db	2521	CAGCTCGAGTGCATGTGGTGCAGTACATGAACAGTGTGTGACCTCCATGCTTACGT	2580
QY	2895	GGCTCTCTCCCTTTGGCCAGTGTATGGAATGTATACGATGAGCAGCTGCCCCCTGA	2954
Db	2581	GGCTCTCTCCCTTTGGCCAGTGTATGGAATGTATACGATGAGCAGCTGCCCCCTGA	2640
QY	2955	AAATGTTTCCAGTCTACTCTGCTGAGTCAATGCTTGGAGCAACAGGCTGTGCTGGTG	3014
Db	2641	AAATGTTTCCAGTCTACTCTGCTGAGTCAATGCTTGGAGCAACAGGCTGTGCTGGTG	2700
QY	3015	TACTGATCCCAAGCAATCTGCAAGGAAATGATAGAGGGTTCCTATAAGACCACT	3074
Db	2701	TACTGATCCCAAGCAATCTGCAAGGAAATGATAGAGGGTTCCTATAAGACCACT	2760
QY	3075	GAGATGCTTGCAGAGCCCTACAGGAAATTTCTATCCAGAGCCCTGCTCAATTCAG	3134
Db	2761	GAGATGCTTGCAGAGCCCTACAGGAAATTTCTATCCAGAGCCCTGCTCAATTCAG	2820
QY	3135	CATGTGCTTAGAGGACAGACATCAACTGTGCTTTTCAATTCAGTGTCCAGCTTGCCAA	3194
Db	2821	CATGTGCTTAGAGGACAGACATCAACTGTGCTTTTCAATTCAGTGTCCAGCTTGCCAA	2880
QY	3195	CAACGGCCACAGTAATGATCAATCAGACATCTGTGAGAGTGTGAGAACCTGACCC	3254
Db	2881	CAACGGCCACAGTAATGATCAATCAGACATCTGTGAGAGTGTGAGAACCTGACCC	2940
QY	3255	AGGCAAGCACTGCGAGACTGCAATCTGCTTCTACCGTCTATCCCAATGAGGAGAA	3314
Db	2941	AGGCAAGCACTGCGAGACTGCAATCTGCTTCTACCGTCTATCCCAATGAGGAGAA	3000
QY	3315	ATGTGAGCCATGCAAGTGCATGGGACGCGTCTGTGTGCAACCAACACGCGCAAGTG	3374
Db	3001	ATGTGAGCCATGCAAGTGCATGGGACGCGTCTGTGTGCAACCAACACGCGCAAGTG	3060
QY	3375	CTTCTGACCAACAGGCGGTCAAGGGGACGAGTGCAGTGTGTGAGGTAGAAATCG	3434
Db	3061	CTTCTGACCAACAGGCGGTCAAGGGGACGAGTGCAGTGTGTGAGGTAGAAATCG	3120
QY	3435	ATACCAAGAAACCCCTCTCAGAGGAACATGTTATTATATCTCTTCTATTGACTATCAGTT	3494
Db	3121	ATACCAAGAAACCCCTCTCAGAGGAACATGTTATTATATCTCTTCTATTGACTATCAGTT	3180
QY	3495	CACCTTTAGTCTATCCAGGAAGATGCTGCTTATTACAGCTATCAATTTTGGCTAC	3554
Db	3181	CACCTTTAGTCTATCCAGGAAGATGCTGCTTATTACAGCTATCAATTTTGGCTAC	3240
QY	3555	TCCTGACGAACAAACAGGATTTGGACATGTTTCAATCAATGCTCCAGAAATTTCAACCT	3614
Db	3241	TCCTGATGAAACAAACAGGATTTGGACATGTTTCAATCAATGCTCCAGAAATTTCAACCT	3300
QY	3615	CAACATCACCTGGGTGCGAGTTTCTCAGCTGGAACCCAGGCTGAGAGAGATGCTCTGT	3674
Db	3301	CAACATCACCTGGGTGCGAGTTTCTCAGCTGGAACCCAGGCTGAGAGAGATGCTCTGT	3360
QY	3675	TGTTTCAAAACCAACATTAAGGAGTACAAAGATGTTTCTTAATGAGAGTTTGATTT	3734
Db	3361	TGTTTCAAAACCAACATTAAGGAGTACAAAGATGTTTCTTAATGAGAGTTTGATTT	3420
QY	3735	TGCGAACCAACCAATATACATTTCTTTTGTATGTTCAGTAATTTCACTGCGCCCAACAA	3794
Db	3421	TGCGAACCAACCAATATACATTTCTTTTGTATGTTCAGTAATTTCACTGCGCCCAACAA	3480
QY	3795	AATTCAGGT	3803

Db

3481

AATTCAGAT

3489

RESULT 4

EX440935

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

EX440935

EX440935

CSODP012YL19

5-PRIME, mRNA sequence.

EX440935

EX440935.1

GI:30781889

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 10212.f For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CSODP012CF10QP1&cluster=10212.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/Invitrogen Corporation 1600

Paradise Avenue Genoscope sequence ID : CSODP012CF10QP1.

Location/Qualifiers

1. 1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CSODP012YL19"

/tissue type="FETAL BRAIN"

/dev\_stage="fetal"

/clone\_lib="Homo sapiens FETAL BRAIN"

/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-strand cDNA was digested with Not I and

cloned into the Not I and EcoRV sites of the pCMVSPORT 6

vector. Library was not normalized."

FEATURES

source

Query Match

Best Local Similarity

Matches

962;

Conservative

4;

Mismatches

8;

Indels

4;

Gaps

3;

QY

2337

GTGCATTTGCCCTGCCGAAATATCTGTGGCATTTGGTGGCAATGCTTAAATTTGTTGTAGGAACCA

2396

Db

57

GTGCATTTGCCCTGCCGAAATATCTGTGGCATTTGGTGGCAATGCTTAAATTTGTTGTAGGAACCA

116

QY

2397

TTTGAAATTTACTACTGCCAAGGAGAAATTATGCAATGCTTAAATTTGTTGTAGGAACCA

2456

Db

117

TTTGAAATTTACTACTGCCAAGGAGAAATTATGCAATGCTTAAATTTGTTGTAGGAACCA

176

QY

2457

CAATGCCCTTTGGCTTCTCTTACACCCAGAGAGGTAGAAATTTGCTCTTAAGCAGCT

2516

Db

177

CAATGCCCTTTGGCTTCTCTTACACCCAGAGGTAGAAATTTGCTCTTAAGCAGCT

236

QY

2517

GCGAATATGAGTCACTCTCAGAGCATGTCACAGCTTACCTTAAACCCCATGGGTGCGCCT

2576

Db

237

GCGAATATGAGTCACTCTCAGAGCATGTCACAGCTTACCTTAAACCCCATGGGTGCGCCT

296

QY

2577

TGGAAGATCAATGTGCTTCTTCTGCTGGAGATATGTCCTCCCATTTCAATAGTTT

2636

Db

297

TGGAAGATCAATGTGCTTCTTCTGCTGGAGATATGTCCTCCCATTTCAATAGTTT

356

QY

2637

ACTACATGGATGCCCTCTGAGCCCATGATGCTGGATTCTGTGGAATTTTATCAGAAC

2696

```

Db 357 ACTACAGTGGATGCGCTCTGAGCCAGTGAATCTGTGAAATTTATCAGAAC 416
QY 2697 CAGTACTCGGGAGTGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTGTGA 2755
Db 417 CAGTACTCGGGAGTGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTGTGA 476
QY 2757 AAGGCTCGAACCACAGTCTAAGCAGTCCGGACACCATGCTGCTTGAGACAGCATG 2816
Db 477 AAGGCTCGAACCACAGTCTAAGCAGTCCGGACACCATGCTGCTTGAGACAGCATG 536
QY 2817 TGGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGCACCAACATCAAGCAGTGTGT 2876
Db 537 TGGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGCACCAACATCAAGCAGTGTGT 596
QY 2877 GGACTCCAAATGCTATGCGCTCTCTCCCTTTGGCCAGTGTATGGAATGTATACAT 2936
Db 597 GGACTCCAAATGCTATGCGCTCTCTCCCTTTGGCCAGTGTATGGAATGTATACAT 656
QY 2937 GAGCACTGCCCCCTGAAATTTTCAGGCTACTGTACTGTACTGTCTTGGAGCA 2996
Db 657 GAGCACTGCCCCCTGAAATTTTCAGGCTACTGTACTGTACTGTCTTGGAGCA 716
QY 2997 ACCAGCTGTGCTGTGTACTGATCCAGCAATATCTGGCAAGGAAATCATAGAGG 3056
Db 717 ACCAGCTGTGCTGTGTACTGATCCAGCAATATCTGGCAAGGAAATCATAGAGG 776
QY 3057 TTCCTATAAGACACAGTGAAGATGCTTCGACGCCCTACAGGAATTTCTATCCACA 3116
Db 777 TTCCTATAAGACACAGTGAAGATGCTTCGACGCCCTACAGGAATTTCTATCCACA 836
QY 3117 GCCCTGTCTCAATCCAGCATGTCTAGAGACAGCAGATACCACTGCTCTTCAATCA 3176
Db 837 GCCCTGTCTCAATCCAGCATGTCTAGAGACAGCAGATACCACTGCTCTTCAATCA 896
QY 3177 CTGTCCAGTTCGCAATGCAAGGCCACAGTAATGCAATCAATCAGACATCTGTGAA 3236
Db 897 CTGTCCAGTTCGCAATGCAAGGCCACAGTAATGCAATCAATCAGACATCTGTGAA 956
QY 3237 GTGTGAGAACTTGACCAAGCAAGCACTGGGAGACCTGCAATATCTGGCTTCTACGGTGA 3296
Db 957 GTGTGAGAACTTGACCAAGCAAGCACTGGGAGACCTGCAATATCTGGCTTCTACGGTGA 1013
QY 3297 TCCCAACCAATGAGGGAA 3314
Db 1014 T-CCCAACCAATGAGGGAA 1030

```

```

RESULT 5
BU506373
LOCUS
DEFINITION
AGENCOURT_10015485 NIH_MGC_94 Mus musculus cdna clone IMAGE:6494441
5', mRNA sequence.
ACCESSION
BU506373
VERSION
EST.
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 964)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cdna Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

```

```

Plate: L14M14050 row: k column: 18
High quality sequence stop: 592.
FEATURES
source
1. 964
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6494441"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
Notes: Organ: eye; Vector: pcwv-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

```

## ORIGIN

```

Query Match 20.3%; Score 775; DB 13; Length 964;
Best Local Similarity 88.9%; Pred. No. 1-2e-135;
Matches 857; Conservative 0; Mismatches 105; Indels 2; Gaps 2;
QY 1996 TGTGTGTGGAAACACAGGGTGTCTCAGTGTATCTCTGGGCGCTGCAACTGATGAACAA 2055
Db 1 TGTGTGTGGAAACACAGGGTGTCTCAGTGTATCTCTGGGCGCTGCAACTGATGAACAA 60
QY 2056 GAAGAAAGTTAAATCAGATGTTTTTCCAAAGACTCTTGACCATCAGAGATGTGAC 2115
Db 61 GCAGAAAGTTAAATCAGAGTGTCTTAAAGAACCTTGACCATCAGAGATGTGAC 120
QY 2116 CAGCACACAGATTTTACAGCTGTACAGCAACACCAATGACTGCCACTGGTGCATGAC 2175
Db 121 CAGCACACAGATTTTACAGCTGTACAGCAACATACCAATGACTGCCACTGGTGCATGAT 180
QY 2176 CATTTGTCTCCCAAGAACACAGCTGTCTCAGAGGCCAGATCTCCATTTTGTAGTATGAG 2235
Db 181 CACTGTGTCTCCCAAGAACACAGCTGTACAGAGGCCAGATCTCCATTTGCAAGTATGAG 240
QY 2236 AATTGCCCCAAGGATAACCCCATCTACTACTGTAAACAAGACAGCTGCAGAGAGCTGT 2295
Db 241 AGTTGCCCCAAGGATAACCCCATCTACTACTGTAAACAAGACAGCTGCAGAGAGCTGT 300
QY 2296 GCCTCGACACAGAACTGCGAGTGGAGGCCCGGAATCAGAGTGCATGTCCTGCCCGAA 2355
Db 301 GCCTCGACACAGAACTGCGAGTGGAGGCCCGGAATCAGAGTGCATGTCCTGCCCGAA 360
QY 2356 AATATCTGTGGCATTGGCTGGCATTTGTTGAAACTCATGTTGAAATTAATCACTGTC 2415
Db 361 AATATCTGTGGCATTGGCTGGCATTTGTTGAAACTCATGTTGAAATTAATCACTGTC 420
QY 2416 AAGGAGAAATATGACAAATGTTTCTGTAGAACCAACAATGCCCTTTTGTCTTCT 2475
Db 421 AAGGAGAAATATGACAAATGTTTCTGTAGAACCAACAATGCCCTTTTGTCTTCT 480
QY 2476 CTTTACCAACCCAGAGAGAGTGAATTTGCTTTCAGAGCTGCGAATATGAGTCACT 2535
Db 481 CTTTACCAACCCAGAGAGAGTGAATTTGCTTTCAGAGCTGCGAATATGAGTCACT 540
QY 2536 CAGAGCATGTCCAAAGCTCACCTTAAACCCCATGGTGGCTTTCGGAAGATCAATGTGTC 2595
Db 541 CAAAGTATGTCCAAAGCTCACCTTAAACCCCATGGTGGCTTTCGGAAGATCAATGTGTC 600
QY 2596 TACTGTGTCTGGAGAGATGTATGCCCATTTACAAATAGTTTACTACAGTGTGATGCCGTCT 2655
Db 601 TACTGTGTCTGGAGAGATGTATGCCCATTTACAAATAGTTTACTACAGTGTGATGCCGTCT 660
QY 2656 GAGCCAGTGTCTGGAGATTTCTGTGGAAATTTTATCAGAACCCAGTACTC-GGGAGTCAA 2714
Db 661 GAGCCAGTGTCTGGAGATTTCTGTGGAAATTTTATCAGAACCCAGTACTC-GGGAGTCAA 720
QY 2715 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCTGTGAAACCAAG 2774
Db 721 GGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCTGTGAAACCAAG 780

```

```
QY 2775 TGCTAAGCAGTCCGGACACCATGCTGGCTTGAGACAG-CATGTGGAGATTGCACGAG 2833
Db 781 TGCCAGCAGTCCGGACACCATGCTGGCTCGCACANCTTGTGGCCAGTGCACCTAGCA 840
QY 2834 CGAGCTCTGAGTGCATGTGGTGCAGCACATGAAGCAGTGTGAGCTCCCAATGCCCTATG 2893
Db 841 CCAGCTCGAGTGCATGTGGTGCAGCACATGAAGCAGTGTGAGCTCCCAATGCCCTATG 900
QY 2894 TGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCCCTG 2953
Db 901 TGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCACTG 960
QY 2954 AAAA 2957
Db 961 AAAA 964

RESULT 6
Bg678679 933 bp mRNA linear EST 01-MAY-2001
LOCUS 602624477F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749427 5',
DEFINITION mRNA sequence.
ACCESSION Bg678679
VERSION Bg678679.1 GI:13910076
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 933)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10602 row: j column: 20
High quality sequence stop: 762.
FEATURES
Location/Qualifiers
1..933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4749427"
/tissue_type="squamous cell carcinoma"
/lab_host="PH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."

ORIGIN
Query Match 20.0%; Score 763.2; DB 12; Length 933;
Best Local Similarity 97.3%; Pred. No. 1.9e-133;
Matches 798; Conservative 0; Mismatches 18; Indels 4; Gaps 2;

QY 2997 ACCAGGCTGGCTGGTGTACTGATCCAGCAATCTGCGCAAGGGAATGATAGAGGG 3056
Db 2 ACCAGGCTGGCTGGTGTACTGATCCAGCAATCTGCGCAAGGGAATGATAGAGGG 61
QY 3057 TTCCCTATAAGGACCACTGAAGTGCCTTGGCAAGCCCTACAGGAATTTCTATCCACA 3116
Db 62 TTCCCTATAAGGACCACTGAAGTGCCTTGGCAAGCCCTACAGGAATTTCTATCCACA 121
QY 3117 GCCCTGCTCAATTCAGCATGTGTAGAGGACACAGATACAACTGCTTTCATTCA 3176
```

```
Db 122 GCCCTGCTCAATTCAGCATGTGTAGAGGACACAGATACAACTGCTTTCATTCA 181
QY 3177 CTGTCAGGCTTGCATATGCAACGGCCACAGTAAATGATCAATCAAGATCTGTGAGAA 3236
Db 182 CTGTCAGGCTTGCATATGCAACGGCCACAGTAAATGATCAATCAAGATCTGTGAGAA 241
QY 3237 GTGTGAGAACCTGACACAGGGAAGCACTGCGAGACCTGCATATCTTACGCTGA 3296
Db 242 GTGTGAGAACCTGACACAGGGAAGCACTGCGAGACCTGCATATCTTACGCTGA 301
QY 3297 TCCCAACAATGGAGGGAAATGTCAGCCATGCAAGTCAATGGGCAAGCGCTCTGTGCAA 3356
Db 302 TCCCAACAATGGAGGGAAATGTCAGCCATGCAAGTCAATGGGCAAGCGCTCTGTGCAA 361
QY 3357 CACCAACAAGGGAAGTGTCTTGTGACCAAGGCGCTCAAGGGGAGAGTGCAGCT 3416
Db 362 CACCAACAAGGGAAGTGTCTTGTGACCAAGGCGCTCAAGGGGAGAGTGCAGCT 421
QY 3417 ATGTGAGGTAGAAATTCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATATCTCT 3476
Db 422 ATGTGAGGTAGAAATTCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATATCTCT 481
QY 3477 TCTTATGATATCAGTTACCTTTAGTCTATCCAGGAGATGATCGCTATTACAGC 3536
Db 482 TCTTATGATATCAGTTACCTTTAGTCTATCCAGGAGATGATCGCTATTACAGC 541
QY 3537 TATCAATTTTGTGCTACTCTCTGACGAACAAACAGGATTTGACATGTTCAATATGC 3596
Db 542 TATCAATTTTGTGCTACTCTCTGACGAACAAACAGGATTTGACATGTTCAATATGC 601
QY 3597 CTCACAGAAATTTCAACCTCAACATCACCTGGGTGCGAGTTCTCAGCTGGAACCCAGGC 3656
Db 602 CTCACAGAAATTTCAACCTCAACATCACCTGGGTGCGAGTTCTCAGCTGGAACCCAGGC 661
QY 3657 TGGAGAAGATGCTGTTGTTTTCACAAACCAACATAGGACTACAAAGATGTTCTC 3716
Db 662 TGGAGAAGATGCTGTTGTTTTCACAAACCAACATAGGACTACAAAGATGTTCTC 721
QY 3717 TAATGAGAAGTGTGATTTTCGCAACCAACCAATATACCTTTCTTTTATGTCAGTAA 3776
Db 722 TAATGAGAAGTGTGATTTTCGCAACCAACCAATATACCTTTCTGTGTGATGTCAGTAA 780
QY 3777 TTTCACT---GGCCCATCAAAATTCAGGTGCAAACTGAA 3813
Db 781 TTTCACTTTGGGCCCCATCGAAATTCAGGTGCAAAATGGA 820

RESULT 7
Cb520374 840 bp mRNA linear EST 09-JUL-2003
LOCUS UI-M-GIO-cei-1-18-0-UT.r1 NIH BMAP_G10 Mus musculus cDNA clone
DEFINITION IMAGE:6840067 5', mRNA sequence.
ACCESSION Cb520374
VERSION Cb520374.1 GI:29353729
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 840)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mouseefi.html
This clone was contributed by the Brain Molecular Anatomy Project
```









```

QY 2992 GAGCAACCGGCTGTGGTGTACTGATCCAGCAATATCGGCAAGGGAAATGCATA 3051
DB 481 GAGCAACCGGCTGTGGTGTACTGATCCAGCAATATCGGCAAGGGAAATGCATA 540
QY 3052 GAGGGTTCCTATAAGGACCACTGGAAGATGCTTCGCAAGCCCTACAGGAATTTCTAT 3111
DB 541 GAGGGTTCCTATAAGGACCACTGGAAGATGCTTCGCAAGCCCTACAGGAATTTCTAT 600
QY 3112 CCACGCCCCCTCTCAATCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 3171
DB 601 CCACGCCCCCTCTCAATCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660
QY 3172 ATTCACTGTCAGCTTGCCAATG 3194
DB 661 ATTCACTGTCAGCTTGCCAATG 683

RESULT 10
BM783258
LOCUS K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5',
DEFINITION mRNA sequence.
ACCESSION BM783258
VERSION BM783258.1 GI:19131490
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 680)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 31 row: H column: 08
High quality sequence stop: 680.
Location/Qualifiers
1. 680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-31-H08"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/notes="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tabacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo di-selected
mRNA by priming with dr-tailed vector. The dr-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation
method."

FEATURES
source
1. 680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-31-H08"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/notes="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tabacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo di-selected
mRNA by priming with dr-tailed vector. The dr-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation
method."

ORIGIN
Query Match 17.8%; Score 680; DB 12; Length 680;
Best Local Similarity 100.0%; Pred. No. 8.9e-118;

```

```

Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2512 CAGCTCGAATAATCAGATCATCTCAGAGCATGTCGAAGTCACCTTAACCCCATGGTTC 2571
DB 1 CAGCTCGAATAATCAGATCATCTCAGAGCATGTCGAAGTCACCTTAACCCCATGGTTC 60
QY 2572 GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCTCAATTAACAAT 2631
DB 61 GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCTCAATTAACAAT 120
QY 2632 AGTTTACTACAGTGGATGGCTGTGAGCCAGTGTGCTGGATTTCTGTGGAAATTTATCA 2691
DB 121 AGTTTACTACAGTGGATGGCTGTGAGCCAGTGTGCTGGATTTCTGTGGAAATTTATCA 180
QY 2692 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTAGTGTTC 2751
DB 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTAGTGTTC 240
QY 2752 TGTGAAGGCTGCAACCCAGTGTGCTGAAGTGTGCGGACACCATGTGCTTGGAGCA 2811
DB 241 TGTGAAGGCTGCAACCCAGTGTGCTGAAGTGTGCGGACACCATGTGCTTGGAGCA 300
QY 2812 GCATGTGGAGATTGCAACCCAGTGTGCTGAAGTGTGCGGACACCATGTGCTTGGAGCA 2871
DB 301 GCATGTGGAGATTGCAACCCAGTGTGCTGAAGTGTGCGGACACCATGTGCTTGGAGCA 360
QY 2872 TGTGTGCAATCCAAATGCTTATGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2931
DB 361 TGTGTGCAATCCAAATGCTTATGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 420
QY 2932 AGCATGAGCACTGCCCCCTGAAATTTGTTCCAGGCTACTGTACCTGTAGTCAATGCTTG 2991
DB 421 AGCATGAGCACTGCCCCCTGAAATTTGTTCCAGGCTACTGTACCTGTAGTCAATGCTTG 480
QY 2992 GAGCAACCGGCTGTGGTGTGTACTGATCCAGCAATATCGGCAAGGGAAATGCATA 3051
DB 481 GAGCAACCGGCTGTGGTGTGTACTGATCCAGCAATATCGGCAAGGGAAATGCATA 540
QY 3052 GAGGGTTCCTATAAGGACCACTGGAAGATGCTTCGCAAGCCCTACAGGAATTTCTAT 3111
DB 541 GAGGGTTCCTATAAGGACCACTGGAAGATGCTTCGCAAGCCCTACAGGAATTTCTAT 600
QY 3112 CCACGCCCCCTCTCAATCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 3171
DB 601 CCACGCCCCCTCTCAATCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660
QY 3172 ATTCACTGTCAGCTTGCCA 3191
DB 661 ATTCACTGTCAGCTTGCCA 680

RESULT 11
BM783773
LOCUS K-EST0061175 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5',
DEFINITION tRNA sequence.
ACCESSION BM783773
VERSION BM783773.1 GI:19132005
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 680)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

```

Tel: +82-42-860-4470  
 Fax: +82-42-860-4470  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 36 row: D column: 02  
 High quality sequence stop: 680.  
 Location/Qualifiers  
 1. 680  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S5SNU484-36-D02"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Epithelial"  
 /lab\_host="SNU-484"  
 /lab\_host="Top10F"  
 /clone\_lib="S5SNU484"  
 /note="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-tailed mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method."

**FEATURES**  
 source  
 1. 680  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S5SNU484-34-F01"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Epithelial"  
 /lab\_host="SNU-484"  
 /lab\_host="Top10F"  
 /clone\_lib="S5SNU484"  
 /note="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method."

**ORIGIN**  
 Query Match 17.8%; Score 680; DB 12; Length 680;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-118;  
 Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2512 CAGCTGCGAATATGTCAGTCTATCTCAGAGCATGTCACAGCTCAGCTTAAACCCCATGGGTC 2571  
 DB 1 CAGCTGCGAATATGTCAGTCTATCTCAGAGCATGTCACAGCTCAGCTTAAACCCCATGGGTC 60

QY 2572 GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAAAT 2631  
 DB 61 GGCCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAAAT 120

QY 2632 AGTTTACTACAGTGCATCCCTCTGAGCCAGTGTGATGCTGATCTGTTGGAAATTTATCA 2691  
 DB 121 AGTTTACTACAGTGCATCCCTCTGAGCCAGTGTGATGCTGATCTGTTGGAAATTTATCA 180

QY 2692 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCCATTTAGTGTGTC 2751  
 DB 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCCATTTAGTGTGTC 240

QY 2752 TGTGAAGGCTGCAAAACACAGTGTGCTAAGCAGTCCGGACACCATGTCCTTGGAGCA 2811  
 DB 241 TGTGAAGGCTGCAAAACACAGTGTGCTAAGCAGTCCGGACACCATGTCCTTGGAGCA 300

QY 2812 GCATGTGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGAGCAACATCAAGCAG 2871  
 DB 301 GCATGTGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGAGCAACATCAAGCAG 360

QY 2872 TGTGTGAGTCCAAATGCGCTATGTGCGCTCTTCCCTTTGGCCAGTGTATGGATGGTAT 2931  
 DB 361 TGTGTGAGTCCAAATGCGCTATGTGCGCTCTTCCCTTTGGCCAGTGTATGGATGGTAT 420

QY 2932 ACGATGAGCACTGCCCTCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCATGCTTG 2991  
 DB 421 ACGATGAGCACTGCCCTCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCATGCTTG 480

QY 2992 GAGCAACAGCGCTGTGGCTGTGTACTGATCCAGCAATCTGCGCAAAAGGAAATGCATA 3051  
 DB 481 GAGCAACAGCGCTGTGGCTGTGTACTGATCCAGCAATCTGCGCAAAAGGAAATGCATA 540

QY 3052 GAGGCTTCTATAAAGGACCACTGAAGATGCCTTCGGAAGCCCTTACAGGAAATTTCTAT 3111

**RESULTS**  
 541 GAGGCTTCTATAAAGGACCACTGAAGATGCCTTCGGAAGCCCTTACAGGAAATTTCTAT 600  
 3112 CCACAGCCCTCTCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 3171  
 601 CCACAGCCCTCTCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660  
 3172 ATTCACTGTCCAGCTTGCCA 3191  
 661 ATTCACTGTCCAGCTTGCCA 680

**LOCUS**  
 BM783718 677 bp mRNA linear EST 05-MAR-2002  
 K-EST0061702 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-F01 5', mRNA sequence.

**ACCESSION**  
 BM783718  
**VERSION**  
 BM783718.1 GI:19131950  
**KEYWORDS**  
 EST.  
**SOURCE**  
 Homo sapiens (human)  
**ORGANISM**  
 Homo sapiens  
**REFERENCE**  
 1 (bases 1 to 677)  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
**AUTHORS**  
 Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.  
**TITLE**  
 21C Frontier Korean EST Project 2001  
**JOURNAL**  
 Unpublished (2002)  
**COMMENT**  
 Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 36 row: F column: 01  
 High quality sequence stop: 677.

**FEATURES**  
 source  
 1. 677  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S5SNU484-34-F01"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Epithelial"  
 /lab\_host="SNU-484"  
 /lab\_host="Top10F"  
 /clone\_lib="S5SNU484"  
 /note="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method."

**ORIGIN**  
 Query Match 17.7%; Score 677; DB 12; Length 677;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-117;  
 Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2512 CAGCTGCGAATATGTCAGTCTATCTCAGAGCATGTCACAGCTCAGCTTAAACCCCATGGGTC 2571  
 DB 1 CAGCTGCGAATATGTCAGTCTATCTCAGAGCATGTCACAGCTCAGCTTAAACCCCATGGGTC 60

FEATURES		Location/Qualifiers	
source		1. 677	
		/organism="Homo sapiens"	
		/mol_type="mRNA"	
		/db_xref="taxon:9606"	
		/clone="S5SNU484-36-H09"	
		/sex="M"	
		/tissue_type="Stomach"	
		/cell_type="Epithelial"	
		/cell_line="SNU-484"	
		/lab_host="Top10F"	
		/clone_lib="S5SNU484"	
		/notes="Organ: Stomach; Vector: pTZ18RPI; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F, by electroporation method."	
ORIGIN		Query Match 17.7%; Score 674.4; DB 12; Length 677;	
		Best Local Similarity 99.7%; Pred. No. 1e-116;	
		Matches 675; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
QY	2512	CAGCTCGGAATTAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC	2571
Db	1	CAGCTCGGAATTAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC	60
QY	2572	GGCTTTGGAGATCAATGTGCTCTGCTGGAGATATGTCCCATTTACAAAT	2631
Db	61	GGCTTTGGAGATCAATGTGCTCTGCTGGAGATATGTCCCATTTACAAAT	120
QY	2632	AGTTTACTACGTGATGCGCTCTGAGCCAGTGTGTTGATCTGTTGGATTTTATCA	2691
Db	121	AGTTTACTACGTGATGCGCTCTGAGCCAGTGTGTTGATCTGTTGGATTTTATCA	180
QY	2692	GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	2751
Db	181	GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	240
QY	2752	TGTGAAGGCTTGAACACACAGTCTTAAGCAGTGGCGGACACCATGTGCCTTGGAGCA	2811
Db	241	TGTGAAGGCTTGAACACACAGTCTTAAGCAGTGGCGGACACCATGTGCCTTGGAGCA	300
QY	2812	GCATGTGGAGATTGCAACCCAGCGGAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAG	2871
Db	301	GCATGTGGAGATTGCAACCCAGCGGAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAG	360
QY	2872	TGTGTGAGCTCAATGCTCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGATGGTAT	2931
Db	361	TGTGTGAGCTCAATGCTCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGATGGTAT	420
QY	2932	ACGATGAGCAGCTGCGCCCTCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCTGCTTG	2991
Db	421	ACGATGAGCAGCTGCGCCCTCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCTGCTTG	480
QY	2992	GAGCAACAGGCTGTGGTGTGTACTCATCCAGCAATATCGGCAAGGAAATGCATA	3051
Db	481	GAGCAACAGGCTGTGGTGTGTACTCATCCAGCAATATCGGCAAGGAAATGCATA	540
QY	3052	GAGGTTTCTTAAAGGACAGTGAAGTGCCTTTCGCAAGCCCTACAGGAAATTTCTAT	3111
Db	541	GAGGTTTCTTAAAGGACAGTGAAGTGCCTTTCGCAAGCCCTACAGGAAATTTCTAT	600
QY	3112	CCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATCAACTGGTCTTTC	3171

QY	2572	GGCTTTGGAGATCAATGTGCTCTACTGCTGCTGGAGATATGTCCCATTTACAAAT	2631
Db	61	GGCTTTGGAGATCAATGTGCTCTACTGCTGCTGGAGATATGTCCCATTTACAAAT	120
QY	2632	AGTTTACTACGTGATGCGCTCTGAGCCAGTGTGTTGATCTGTTGGATTTTATCA	2691
Db	121	AGTTTACTACGTGATGCGCTCTGAGCCAGTGTGTTGATCTGTTGGATTTTATCA	180
QY	2692	GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	2751
Db	181	GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	240
QY	2752	TGTGAAGGCTTGAACACACAGTCTTAAGCAGTGGCGGACACCATGTGCCTTGGAGCA	2811
Db	241	TGTGAAGGCTTGAACACACAGTCTTAAGCAGTGGCGGACACCATGTGCCTTGGAGCA	300
QY	2812	GCATGTGGAGATTGCAACCCAGCGGAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAG	2871
Db	301	GCATGTGGAGATTGCAACCCAGCGGAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAG	360
QY	2872	TGTGTGAGCTCAATGCTCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGATGGTAT	2931
Db	361	TGTGTGAGCTCAATGCTCTATGTGGCTCTTCCCTTTTGGCCAGTGTATGGATGGTAT	420
QY	2932	ACGATGAGCAGCTGCGCCCTCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCTGCTTG	2991
Db	421	ACGATGAGCAGCTGCGCCCTCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCTGCTTG	480
QY	2992	GAGCAACAGGCTGTGGTGTGTACTCATCCAGCAATATCGGCAAGGAAATGCATA	3051
Db	481	GAGCAACAGGCTGTGGTGTGTACTCATCCAGCAATATCGGCAAGGAAATGCATA	540
QY	3052	GAGGTTTCTTAAAGGACAGTGAAGTGCCTTTCGCAAGCCCTACAGGAAATTTCTAT	3111
Db	541	GAGGTTTCTTAAAGGACAGTGAAGTGCCTTTCGCAAGCCCTACAGGAAATTTCTAT	600
QY	3112	CCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATCAACTGGTCTTTC	3171
Db	601	CCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATCAACTGGTCTTTC	660
QY	3172	ATTCACTGTCCAGCTTG 3188	
Db	661	ATTCACTGTCCAGCTTG 677	

RESULT 13	BM783821	677 bp	mRNA	linear	EST 05-MAR-2002				
LOCUS	K-EST0061830	S5SNU484	Homo sapiens	cDNA clone	S5SNU484-36-H09 5',				
DEFINITION	mRNA sequence.								
ACCESSION	BM783821								
VERSION	BM783821.1	GI:19132053							
KEYWORDS	EST.								
SOURCE	Homo sapiens (human)								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.								
AUTHORS	1 (bases 1 to 677) Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.								
TITLE	21C Frontier Korean EST Project 2001								
JOURNAL	Unpublished (2002)								
COMMENT	Contact: Kim YS Genome Research Center Korea Research Institute of Bioscience & Biotechnology 52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea Tel: +82-42-860-4470 Fax: +82-42-860-4409 Email: yongsung@mail.kribb.re.kr Plate: 36 row: H column: 09 High quality sequence stop: 677.								

```

Db 601 CCAGAGCCCTGCTCAATCCAGATGTGTCTAGAGGACAGACAGATACATGCTCTTTC 660
QY 3172 ATTCACTGTCAGCTTG 3188
Db 661 ATTCACTGTCAGCTTG 677

RESULT 14
BM783685
LOCUS K-EST0061659 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-B06 5',
DEFINITION mRNA sequence.
ACCESSION BM783685
VERSION BM783685.1 GI:19131917
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 673)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
Unpublished (2002)
Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-gong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34 row: B column: 06
High quality sequence stop: 673.
FEATURES
location/Qualifiers
1..673
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-34-B06"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_host="Top10P"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ19R1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P, by electroporation
method."

ORIGIN
Query Match 17.6%; Score 673; DB 12; Length 673;
Best Local Similarity 100.0%; Pred. No. 1.9e-116; Mismatches 0; Indels 0; Gaps 0;
Matches 673; Conservative 0;

QY 2512 CAGCTCGGAATATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC 2571
Db 1 CAGCTCGGAATATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTC 60

QY 2572 GGCCTTCGGAAGATCAATGTGTCCTACTGCTGGAGAGATATGTCCCATTTACAAAT 2631
Db 61 GGCCTTCGGAAGATCAATGTGTCCTACTGCTGGAGAGATATGTCCCATTTACAAAT 120

```

```

QY 2632 AGTTTACTACAGTGGATGCCGTCTGAGCCCAAGTGAAGTCTGTGGATTCGTGTGAAATTTATCA 2691
Db 121 AGTTTACTACAGTGGATGCCGTCTGAGCCCAAGTGAAGTCTGTGTGAAATTTATCA 180
QY 2692 GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGTGATCAACCCACTCAATGGTAGTTC 2751
Db 181 GAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGTGATCAACCCACTCAATGGTAGTTC 240
QY 2752 TGTGAAGGGCTCTCAACACACAGTGTAAAGCAGTGGCGGACACCATGTGCTTGGAGACA 2811
Db 241 TGTGAAGGGCTCTCAACACACAGTGTAAAGCAGTGGCGGACACCATGTGCTTGGAGACA 300
QY 2812 GCATGTGGAGATTGCACAGCGGCGCTCTGAGTGCATGTGGTGGACGACACATGAAGCAG 2871
Db 301 GCATGTGGAGATTGCACAGCGGCGCTCTGAGTGCATGTGGTGGACGACACATGAAGCAG 360
QY 2872 TGTGTGCACTCCCAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTAT 2931
Db 361 TGTGTGCACTCCCAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTAT 420
QY 2932 ACGATGAGCACCTGCCCCCTGAAATTTGTTCAGGCTACTGTACCTGTAGTCAATGCTTG 2991
Db 421 ACGATGAGCACCTGCCCCCTGAAATTTGTTCAGGCTACTGTACCTGTAGTCAATGCTTG 480
QY 2992 GAGCAACCGGCTGTGGCTGGTGTACTGATCCGAGCAATCTGGCAAGGGAATGCATA 3051
Db 481 GAGCAACCGGCTGTGGCTGGTGTACTGATCCGAGCAATCTGGCAAGGGAATGCATA 540
QY 3052 GAGGTTTCCTTATAAGGACCAAGTGAAGATGCCCTTCGCAAGCCCTCAGAGAAATTTCTAT 3111
Db 541 GAGGTTTCCTTATAAGGACCAAGTGAAGATGCCCTTCGCAAGCCCTCAGAGAAATTTCTAT 600
QY 3112 CCACAGCCCTGTCTCAATTCACGATGTGTCTAGAGACAGCAGATACAACTGGTCTTTC 3171
Db 601 CCACAGCCCTGTCTCAATTCACGATGTGTCTAGAGACAGCAGATACAACTGGTCTTTC 660
QY 3172 ATTCACGTGTCAG 3184
Db 661 ATTCACGTGTCAG 673

RESULT 15
BX431593 962 bp mRNA linear EST 15-MAY-2003
LOCUS BX431593
DEFINITION BX431593 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CSODP012YL19 5-PRIME, mRNA sequence.
ACCESSION BX431593
VERSION BX431593.1 GI:30781048
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 962)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10212.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAG017ZB02_CS01546_1&cluster=10212.f.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAG017ZB02_CS01546_1.
FEATURES
Location/Qualifiers
1..962
/organism="Homo sapiens"
/mol_type="mRNA"

```

```
/db xref="taxon:9606"  
/clone="CS0DP012YL19"  
/tissue_type="FETAL BRAIN"  
/dev_stage="fetal"  
/clone_lib="Homo sapiens FETAL BRAIN"  
/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-strand cDNA was digested with Not I and  
cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
vector. Library was not normalized."
```

## ORIGIN

```
Query Match 17.5%; Score 669.8; DB 13; Length 962;  
Best Local Similarity 98.4%; Pred. No. 8.6e-116;  
Matches 697; Conservative 0; Mismatches 8; Indels 3; Gaps 2;  
  
QY 3113 CACAGCCCGTCTCAATCCAGCATGTGTCTAGAGACAGCAGATA-CAACTGGTCTTTC 3171  
Db |||||||  
7 CACAGCCCGTCTCT--AATTCAGATGTGTCTAGAGACAGCAGATACCACCTGGTCTTTC 64  
  
QY 3172 ATTCACTGTCCAGCTTGCCCAATGCAACGCCACACAGTAAATGCATCAATCAGAGCATCTGT 3231  
Db |||||||  
65 ATTCACTGTCCAGCTTGCCCAATGCAACGCCACACAGTAAATGCATCAATCAGAGCATCTGT 124  
  
QY 3232 GAGAAGTGTGAGAACTGTACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 3291  
Db |||||||  
125 GAGAAGTGTGAGAACTGTACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 184  
  
QY 3292 GGTGATCCCAACCAATGGAGGGAATGTTCAGCCATGCAATGCAATGGGCAAGCGTCTCTG 3351  
Db |||||||  
185 GGTGATCCCAACCAATGGAGGGAATGTTCAGCCATGCAATGCAATGGGCAAGCGTCTCTG 244  
  
QY 3352 TGCACACCAACACACGGGCAAGTGTCTTCGACCAACCAAGGGCGTCAAGGGGGAAGAGTGC 3411  
Db |||||||  
245 TGCACACCAACACACGGGCAAGTGTCTTCGACCAACCAAGGGCGTCAAGGGGGAAGAGTGC 304  
  
QY 3412 CAGCTATGTGAGGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTAT 3471  
Db |||||||  
305 CAGCTATGTGAGGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTAT 364  
  
QY 3472 ACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTTATCCAGGAAGATGATCGCTATTAC 3531  
Db |||||||  
365 ACTCTTCTTATTGACTATCAGTTCACCTTTAGTCTTATCCAGGAAGATGATCGCTATTAC 424  
  
QY 3532 ACAGCTATCAATTTTGTGGTACTCTCTACGAAACAAACAGGGATTTGGACATGTTTCATC 3591  
Db |||||||  
425 ACAGCTATCAATTTTGTGGTACTCTCTACGAAACAAACAGGGATTTGGACATGTTTCATC 484  
  
QY 3592 AATGCTCCAGAAATTCAACCTCAACATCACCTGGGCTGCAGTTTCTCAGCTGGAACC 3651  
Db |||||||  
485 AATGCTCCAGAAATTCAACCTCAACATCACCTGGGCTGCAGTTTCTCAGCTGGAACC 544  
  
QY 3652 CAGGTGGAGAGAGATGCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGT 3711  
Db |||||||  
545 CAGGTGGAGAGAGATGCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGT 604  
  
QY 3712 TTCTCTAATGAGAGTTGATTTTCGCAACCAACCAATATCACTTCTTGTGTTTATGTC 3771  
Db |||||||  
605 TTCTCTAATGAGAGTTGATTTTCGCAACCAACCAATATCACTTCTTGTGTTTATGTC 664  
  
QY 3772 AGTAATTTCACTGGCCCATCAAAATTCAGGTGCAAACTGAACATGA 3819  
Db |||||||  
665 AGTAATTTCACTGGCCCATCAAAATTCAGGTGCAAACTGAACATGA 712
```

Search completed: March 5, 2004, 12:08:03  
Job time : 6054.99 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 4, 2004, 16:06:34 ; Search time 899.913 Seconds  
(without alignments)  
18028.269 Million cell updates/sec

Title: US-09-787-097-19  
Perfect score: 3819  
Sequence: 1 atggtgcccagcgccggc.....aggtgcaactgaacaatga 3819

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : N\_Geneseq\_29Jan04.\*
- 1: Geneseqn1980s.\*
  - 2: Geneseqn1990s.\*
  - 3: Geneseqn2000s.\*
  - 4: Geneseqn2001as.\*
  - 5: Geneseqn2001bs.\*
  - 6: Geneseqn2002s.\*
  - 7: Geneseqn2003as.\*
  - 8: Geneseqn2003bs.\*
  - 9: Geneseqn2003cs.\*
  - 10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3819	100.0	3819	3	Aaz52277 Human sol
2	3801.4	99.5	4290	3	Aaz52276 Human mem
3	3801.4	99.5	8561	5	Aaz72659 DNA encod
4	3713.8	97.2	4072	3	Aaz91921 Human mah
5	3696.2	96.8	8590	3	Aaz91920 Human mah
6	3492.4	91.4	4068	3	Aaz52275 Human mem
7	3487	91.3	3597	3	Aaz52274 Human sol
8	3487	91.3	3597	5	Aaz52267 DNA encod
9	3487	91.3	3597	9	ADB70997 Human att
10	3114.8	81.6	8739	9	ADB53249 Primary r
11	2338.2	76.9	8827	3	Aaz91916 Wild type
12	2810.2	73.6	3490	5	Aaz72660 DNA encod
13	2248.2	58.9	2625	3	Aaz91922 Human mah
14	1521.8	39.8	2419	3	Aaz91917 Murine mah
15	1341.6	35.1	6373	3	Aaz91919 Human mah
16	1233.6	32.3	4134	7	Abz58710 Murine ho
17	1228.2	32.2	4140	7	Abz58709 Human hom
18	1135	29.7	3609	6	Abz53434 Human cdn
19	873.2	22.9	3165	5	Aaz78781 DNA encod
20	873.2	22.9	3313	5	Aaz76294 DNA encod
21	819.4	21.5	1012	2	Aax84568 Biorhythm
22	640.2	16.8	1051	3	Aaz91918 Murine mah
23	574	15.0	6517	4	Aah98748 Human EST

ALIGNMENTS

RESULT 1

Aaz52277  
ID AAZ52277 standard; CDNA; 3819 BP.  
XX AC AAZ52277;  
XX AC  
DT 18-JUL-2000 (first entry)  
DE Human soluble attractin-2 CDNA.  
XX Human; soluble attractin-2; immune response; macrophage; monocyte;  
KW T cell; immunostimulant; immunosuppressed patient; cancer;  
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
FH CDS 1..3819  
FT /\*tag= a  
FT /product= "Soluble attractin-2"  
XX  
PN WO200015651-A1.  
XX 23-MAR-2000.  
XX  
XX 14-SEP-1999; 99WO-US020948.  
XX  
XX 14-SEP-1998; 98US-0100137P.  
XX (DAND ) DANA FARBER CANCER INST INC.  
XX Duke-Cohan JS, Schlossman SF;  
XX WPI; 2000-271373/23.  
XX P-PSDB; AAY70692.  
XX Isolated nucleic acids encoding human attractin polypeptides useful for enhancing immune responses.  
XX Claim 3; Fig 12; 120pp; English.  
XX The patent discloses four forms of human attractin polypeptides which enhance immune response by promoting macrophage and monocyte spreading in the presence of T cells. These include soluble attractin-1 and -2 and membrane attractin-1 and -2. These various forms of attractin are encoded by alternatively spliced mRNA molecule transcribed from a single gene.

Aas86766 DNA encod  
Aas72658 DNA encod  
Abx34730 Human mad  
Aal54064 CDNA of h  
Aas02409 Human sec  
Aal14686 Human bre  
Aaa45505 Human sec  
Abq52394 Oligonuc  
Abq52395 Oligonuc  
Abq52396 Oligonuc  
Aal23548 Human bre  
Aal39784 Human NS  
Aal43425 Probe #12  
Abs11523 Human gen  
Abk38928 CDNA enco  
Acal1257 Human lun  
Aca02443 Lung canc  
Abz58714 Upregulat  
Abz72040 Gene 216  
Abx74891 BAC1098L2  
Ab132196 Human imm  
Abq52396 Oligonuc  
Abq52397 Oligonuc





1921 CTGGTATTCACCTCGGAACAGTGTGATCGGCATCGGAGTGAAGCGCTTGTATGAGCA 1980  
1981 GGACCTGGTATTCAGTGTGTGTGGAACACAGGGTGTCTCAGTGTATCTCGTGGGCGCTG 2040  
1981 GGACCTGGTATTCAGTGTGTGTGGAACACAGGGTGTCTCAGTGTATCTCGTGGGCGCTG 2040  
2041 GCAACTGATGAACAGGAAGAAAGTTAAATCAGATGTTTCCAAAAGAACTCTTGAC 2100  
2041 GCAACTGATGAACAGGAAGAAAGTTAAATCAGATGTTTCCAAAAGAACTCTTGAC 2100  
2101 CATGACAGATGTGACACAGCACACAGATGTTTACAGCTGTACAGCAACACCAATGACTGC 2160  
2101 CATGACAGATGTGACACAGCACACAGATGTTTACAGCTGTACAGCAACACCAATGACTGC 2160  
2161 CACTGGTGCATGACCAATGTGTCCCAAGAAACACAGCTGTCTCAGAGGCCAGATCTCC 2220  
2161 CACTGGTGCATGACCAATGTGTCCCAAGAAACACAGCTGTCTCAGAGGCCAGATCTCC 2220  
2221 ATTTTATAGGTATGAGAAATGCCCCAAGATTAACCCCATGTACTCTGTAAACAGAGAAC 2280  
2221 ATTTTATAGGTATGAGAAATGCCCCAAGATTAACCCCATGTACTCTGTAAACAGAGAAC 2280  
2281 AGCTGCAGGAGCTGTGCCCTGGACAGAACTGCAAGTGGAGGCCCGGAATCAGGAGTGC 2340  
2281 AGCTGCAGGAGCTGTGCCCTGGACAGAACTGCAAGTGGAGGCCCGGAATCAGGAGTGC 2340  
2341 ATTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTGGAACTCATGTTTG 2400  
2341 ATTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTGGAACTCATGTTTG 2400  
2401 AAAATTACTACTGCCAAGGAGAAATATGACAATGCTAAATTTGTTCTGTAGGAACCACAAT 2460  
2401 AAAATTACTACTGCCAAGGAGAAATATGACAATGCTAAATTTGTTCTGTAGGAACCACAAT 2460  
2461 GCCCTTTGGCTTCTTTAACCCAGAGAGAGTGAATTTGTCCTTAAGCAGCTGGA 2520  
2461 GCCCTTTGGCTTCTTTAACCCAGAGAGAGTGAATTTGTCCTTAAGCAGCTGGA 2520  
2521 ATATGCACTGATCTCAGAGATGTCACAGTCCACCTTAACCCCATGGCTGGGCTTCGG 2580  
2521 ATATGCACTGATCTCAGAGATGTCACAGTCCACCTTAACCCCATGGCTGGGCTTCGG 2580  
2581 AAGATCAATGTCTCTACTGTGTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTA 2640  
2581 AAGATCAATGTCTCTACTGTGTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTA 2640  
2641 CAGTGGATGCCGTCTGAGCCAGTGTGATGCTGGATTCCTGGAAATTTATCAGAACCCAGT 2700  
2641 CAGTGGATGCCGTCTGAGCCAGTGTGATGCTGGATTCCTGGAAATTTATCAGAACCCAGT 2700  
2701 ACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTGTGTCTGTCAAAGG 2760  
2701 ACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTGTGTCTGTCAAAGG 2760  
2761 CTTGCAAAACCAAGTCTAAGCAGTGCAGGACACCATGTGCTTGAGGACAGCATGTGGA 2820  
2761 CTTGCAAAACCAAGTCTAAGCAGTGCAGGACACCATGTGCTTGAGGACAGCATGTGGA 2820  
2821 GATTCACACAGGCGAGCTCTCAGTGTGATGTGTCAGCAACATGAAGCAGTGTGTGAC 2880  
2821 GATTCACACAGGCGAGCTCTCAGTGTGATGTGTCAGCAACATGAAGCAGTGTGTGAC 2880  
2881 TCCAAATGCCATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGTATACAGTATGAC 2940  
2881 TCCAAATGCCATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGTATACAGTATGAC 2940  
2941 ACTGCCCCCTGAAATTTGTCAGCTACTGTACTGTAGTCATGCTGTGGAGCAACA 3000  
2941 ACTGCCCCCTGAAATTTGTCAGCTACTGTACTGTAGTCATGCTGTGGAGCAACA 3000  
3001 GGCTGTGGCTGTGTACTGATCCCAAGCAATATCGGCAAGGGAATGCATAGAGGTTCC 3060

3001 GGCTGTGGCTGTGTACTGATCCCAAGCAATACTGGCAAAAGGAAATGCATAGAGGTTCC 3060  
3061 TATTAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGAGCC 3120  
3061 TATTAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGAGCC 3120  
3121 CTGCTCAATTCAGCATGTCTAGAGGACAGCAGATACAACCTGGTCTTTCACTCACTGT 3180  
3121 CTGCTCAATTCAGCATGTCTAGAGGACAGCAGATACAACCTGGTCTTTCACTCACTGT 3180  
3181 CAGCTTGCATGCAACGCGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGT 3240  
3181 CAGCTTGCATGCAACGCGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGT 3240  
3241 GAGAACTGCACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTTCTACCGTGATCC 3300  
3241 GAGAACTGCACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTTCTACCGTGATCC 3300  
3301 ACCAATGGAGGAAATGTGAGCATGCAAGTGCATGGCAATGGGACGCTCTGTGCAACACC 3360  
3301 ACCAATGGAGGAAATGTGAGCATGCAAGTGCATGGCAATGGGACGCTCTGTGCAACACC 3360  
3361 AACACGGGCAAGTGTCTGACACCAAGGCGGTCAAGGGGACAGTGCAGCTATGT 3420  
3361 AACACGGGCAAGTGTCTGACACCAAGGCGGTCAAGGGGACAGTGCAGCTATGT 3420  
3421 GAGGTAGAAAATCGATPACCAAGGAAACCTCTCAGAGGAACATGTTATTATCTCTCTT 3480  
3421 GAGGTAGAAAATCGATPACCAAGGAAACCTCTCAGAGGAACATGTTATTATCTCTCTT 3480  
3481 ATTGACTATCAGTTCACTTTAGTCTATCCAGGAGATGATCGCTATTACACAGCTATC 3540  
3481 ATTGACTATCAGTTCACTTTAGTCTATCCAGGAGATGATCGCTATTACACAGCTATC 3540  
3541 AATTTTGTGGCTACTCCTGACGAAACAAACAGGGATTTGGACATGTTCAATATGCTCTC 3600  
3541 AATTTTGTGGCTACTCCTGACGAAACAAACAGGGATTTGGACATGTTCAATATGCTCTC 3600  
3601 AAGAAATTCACCTCAACATCAGTGGCTGCCAGTTCTCAGCTGGAAACCCAGGCTGA 3660  
3601 AAGAAATTCACCTCAACATCAGTGGCTGCCAGTTCTCAGCTGGAAACCCAGGCTGA 3660  
3661 GAGAGATGCTGCTGTTTCAAAAACCAACATTAAGGAGTACAAAAGATAGTTTCTCTAAT 3720  
3661 GAGAGATGCTGCTGTTTCAAAAACCAACATTAAGGAGTACAAAAGATAGTTTCTCTAAT 3720  
3721 GAGAAGTTGATTTTCGCAACCCCAATATCATTCTTTGTTTATGTCAGTAATTC 3780  
3721 GAGAAGTTGATTTTCGCAACCCCAATATCATTCTTTGTTTATGTCAGTAATTC 3780  
3781 ACCTGGCCCATCAAAATTCAGGTGCAACTGACAAATGA 3819  
3781 ACCTGGCCCATCAAAATTCAGGTGCAACTGACAAATGA 3819

RESULT 2  
AAZ52276

ID AAZ52276 standard; cDNA; 4290 BP.

XX AAZ52276;

AC AC (first entry)

DT 18-JUL-2000

XX Human membrane attractin-2 cDNA.

Human; membrane attractin-2; immune response; macrophage; monocyte;  
T cell; immunostimulant; immunosuppressed patient; cancer;  
immunodeficiency syndrome; transplant; autoimmune disease; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1.4290

```

FT      /tag= a
XX      /product= "Membrane attractin-2"
XX      WO200015651-A1.
XX      PN
XX      PD      23-MAR-2000.
XX      PF      14-SEP-1999; 99WO-US020948.
XX      PR      14-SEP-1998; 98US-0100137P.
XX      PA      (DAND ) DANA FARBER CANCER INST INC.
XX      PI      Duke-Cohan JS, Schlossman SF;
XX      WPI; 2000-271373/23.
XX      P-PSDB; AAY70691.
XX      Isolated nucleic acids encoding human attractin polypeptides useful for
XX      enhancing immune responses.
XX      Claim 3; Fig 14; 120pp; English.
XX      The patent discloses four forms of human attractin polypeptides which
XX      enhance immune response by promoting macrophage and monocyte spreading in
XX      the presence of T cells. These include soluble attractin-1 and -2 and
XX      membrane attractin-1 and -2. These various forms of attractin are encoded
XX      by alternatively spliced mRNA molecule transcribed from a single gene.
XX      The present sequence is a cDNA encoding human membrane attractin-2.
XX      Membrane attractin differs from soluble attractin in having a
XX      transmembrane domain and a cytoplasmic domain at the C-terminal.
XX      Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
XX      the N-terminal. This sequence can be used to enhance immune response in
XX      immunosuppressed patients such as those undergoing chemo- and radio-
XX      therapy treatment for cancer or those suffering from common variable
XX      immunodeficiency syndrome. The proteins may also be used to screen
XX      modulators (agonists and antagonists) of immune response which may also
XX      be used to regulate immune reactions. Attractin antibodies can be used to
XX      inhibit immune response in transplant recipients or patients afflicted
XX      with autoimmune disease
XX      SQ      Sequence 4290 BP; 1060 A; 1025 C; 1120 G; 1085 T; 0 U; 0 Other;
XX      Query Match      99.5%; Score 3801.4; DB 3; Length 4290;
XX      Best Local Similarity 100.0%; Pred. No. 0;
XX      Matches 3802; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX      1 ATGGTGGCGGAGCGGGCGGCACTGAGGCAAGGCTGAGGAGGAGGAGCGCGGCGACGGCA 60
XX      1 ATGGTGGCGGAGCGGGCGGCACTGAGGCAAGGCTGAGGAGGAGGAGCGCGGCGACGGCA 60
XX      61 GCGCTCGCGGCGAGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
XX      61 GCGCTCGCGGCGAGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
XX      121 CCGGGGCTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
XX      121 CCGGGGCTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
XX      181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
XX      181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
XX      241 GCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
XX      241 GCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300
XX      301 TGTGACCGGCGGCTGTGTCAACGGGGGTGCTGCAACCGTGGCAACGGGCGAGTGGCTGTCG 360
XX      301 TGTGACCGGCGGCTGTGTCAACGGGGGTGCTGCAACCGTGGCAACGGGCGAGTGGCTGTCG 360
XX      361 CCCGCGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGCGGCGGCGGCGGCGGCGGCGGCG 420

```

```

Db      361 CCCGCCGCTGGGTGGCGGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAACTGGA 420
Qy      421 TCCTCTGGGTTTGTGACAGATGACCTGGGAAATATAAATACAAACGAGTGCACGTGG 480
Db      421 TCCTCTGGGTTTGTGACAGATGACCTGGGAAATATAAATACAAACGAGTGCACGTGG 480
Qy      481 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTGCTACAGAG 540
Db      481 CTCATTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTGCTACAGAG 540
Qy      541 TGTAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGCAACGGTAGTTGCT 600
Db      541 TGTAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGCAACGGTAGTTGCT 600
Qy      601 GCATTTAGTGGCTCATTTGTTCTCTGAGAGAGATGGCAATGAGACTCTCCCTGAGGTTGT 660
Db      601 GCATTTAGTGGCTCATTTGTTCTCTGAGAGAGATGGCAATGAGACTCTCCCTGAGGTTGT 660
Qy      661 GCCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db      661 GCCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy      721 TTTAATATTACTTACAGTTTGTGATGTTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Db      721 TTTAATATTACTTACAGTTTGTGATGTTGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
Qy      781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAACCTGGAAGGTGAA 840
Db      781 AAGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAACCTGGAAGGTGAA 840
Qy      841 GCATGTGACATTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db      841 GCATGTGACATTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy      901 TCAAGTCAATGCTCAGAGGATGCTCCTCTCTCAGACTGCGAGGCTGCTGCTGCTGCTGCTGCT 960
Db      901 TCAAGTCAATGCTCAGAGGATGCTCCTCTCTCAGACTGCGAGGCTGCTGCTGCTGCTGCTGCT 960
Qy      961 CTGTGACAGCTAAACAGTCAATTTGGACTCGAGAGGAAATATTTAACTTAAAGTCTCCC 1020
Db      961 CTGTGACAGCTAAACAGTCAATTTGGACTCGAGAGGAAATATTTAACTTAAAGTCTCCC 1020
Qy      1021 AGAGCATCTCAATAAGCTGTGTCAATGAAACATTTATGCGGTGTTGTTGAGGATATATG 1080
Db      1021 AGAGCATCTCAATAAGCTGTGTCAATGAAACATTTATGCGGTGTTGTTGAGGATATATG 1080
Qy      1081 TTCAACCATCTCAGATTTAATCAATGTTTCTAGCGTATGACCTTCTTCTAGGAGTGGCTT 1140
Db      1081 TTCAACCATCTCAGATTTAATCAATGTTTCTAGCGTATGACCTTCTTCTAGGAGTGGCTT 1140
Qy      1141 CCACTTAAACCGTCTGTGAAACATGTTGTTAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1200
Db      1141 CCACTTAAACCGTCTGTGAAACATGTTGTTAGTATGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1200
Qy      1201 AAGGATAAAATTTACATGTTATGAGGAGGAAAAATGATCCAACTGGGAAATGTGACCAATGAG 1260
Db      1201 AAGGATAAAATTTACATGTTATGAGGAGGAAAAATGATCCAACTGGGAAATGTGACCAATGAG 1260
Qy      1261 TTGAGAGTTTTCATATTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db      1261 TTGAGAGTTTTCATATTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Qy      1321 CAGTATGAGTGGTGGGCACTCTGCAACATTTTACACTGAAGAAATGGCGGAGTGGTC 1380
Db      1321 CAGTATGAGTGGTGGGCACTCTGCAACATTTTACACTGAAGAAATGGCGGAGTGGTC 1380
Qy      1381 ATGCTGGTCACTTTGGTCACTGCGCTCTCTATGATATATAAGCAATGTGAGGAATAT 1440
Db      1381 ATGCTGGTCACTTTGGTCACTGCGCTCTCTATGATATATAAGCAATGTGAGGAATAT 1440
Qy      1441 GATTTGGATTAAGAACACATGAGGATATATTACACCCAGGGTGGCTTGTGCAAGGGGT 1500
Db      1441 GATTTGGATTAAGAACACATGAGGATATATTACACCCAGGGTGGCTTGTGCAAGGGGT 1500

```



Db 3661 GAAGAGATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTCTCTAAT 3720  
QY 3721 GAGAAGTTTGATTTTCGCAACACCAATATACACTTTCTTTGTTATGTACAGTAATTC 3780  
Db 3721 GAGAAGTTTGATTTTCGCAACACCAATATATACACTTTCTTTGTTATGTACAGTAATTC 3780  
QY 3781 ACCTGCCCATCAAAATTCAGGT 3803  
Db 3781 ACCTGCCCATCAAAATTCAGAT 3803

RESULT 3  
AAS72659  
ID AAS72659 standard; cDNA; 8561 BP.  
XX AAS72659;  
XX  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #8463.  
XX  
XX Human; Chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
PI  
XX  
XX WPI; 2001-639362/73.  
DR P-PSDB; ABG08472.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 1; SEQ ID NO 8463; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 8561 BP; 2246 A; 1983 C; 2065 G; 2267 T; 0 U; 0 Other;

Query Match 99.5%; Score 3801.4; DB 5; Length 8561;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3802; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGTTGCGCGAGCGCGGCAACTGAGGCAAGCTGAGGAGGAGGACGCGCGGCGAGCGGA 60  
Db 1 ATGTTGCGCGAGCGCGGCAACTGAGGCAAGCTGAGGAGGAGGACGCGCGGCGAGCGGA 60  
QY 61 GCGCTCGCGGCGAGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
Db 61 GCGCTCGCGGCGAGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
QY 121 CCGGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
Db 121 CCGGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
QY 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
Db 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
QY 241 GCGGAGCG 300  
Db 241 GCGGAGCG 300  
QY 301 TGTGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 360  
Db 301 TGTGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 360  
QY 361 CCGCGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420  
Db 361 CCGCGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 420  
QY 421 TCCTTCTGGGTTTGTGACAGATGACCTGGAATTAATAAATAAATAAATAAATAAATA 480  
Db 421 TCCTTCTGGGTTTGTGACAGATGACCTGGAATTAATAAATAAATAAATAAATAAATA 480  
QY 481 CTCATTGAGGACAGCGCAATAGAAATAGAAATAGAAATAGAAATAGAAATAGAAATAG 540  
Db 481 CTCATTGAGGACAGCGCAATAGAAATAGAAATAGAAATAGAAATAGAAATAGAAATAG 540  
QY 541 TGTAGTTGGGACCATTTATATGTTTATGATGGGAGCTCAATTTATGACCGCTAGTGT 600  
Db 541 TGTAGTTGGGACCATTTATATGTTTATGATGGGAGCTCAATTTATGACCGCTAGTGT 600  
QY 601 GCATTTAGTGGGCTCATTTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGTGTT 660  
Db 601 GCATTTAGTGGGCTCATTTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGTGTT 660  
QY 661 GCCACATCAGGTTATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
Db 661 GCCACATCAGGTTATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720  
QY 721 TTTAATATTACTTACAGTTTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 780  
Db 721 TTTAATATTACTTACAGTTTGTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTT 780  
QY 781 AAGATCAGTAATAGCAGGCAAACTGTTGAATGTTGAATGTTGAATGTTGAATGTTGA 840  
Db 781 AAGATCAGTAATAGCAGGCAAACTGTTGAATGTTGAATGTTGAATGTTGAATGTTGA 840  
QY 841 GCATGTGACATTCCTCATCTGTATACAGAACTGTGGTTTTTCTCATCGAGGCAATCTG 900  
Db 841 GCATGTGACATTCCTCATCTGTATACAGAACTGTGGTTTTTCTCATCGAGGCAATCTG 900  
QY 901 TCAAGTGATGTACAGGAGTCTCTGCTTCTCAGACTGCGAGGCTCTCGGAGTTTCAGTT 960  
Db 901 TCAAGTGATGTACAGGAGTCTCTGCTTCTCAGACTGCGAGGCTCTCGGAGTTTCAGTT 960  
QY 961 CTTGTACCAAGCTAACAGTCAATTTTGGACTCGAGAGGAATATTCTAATCTAAAGCTCCC 1020  
Db 961 CTTGTACCAAGCTAACAGTCAATTTTGGACTCGAGAGGAATATTCTAATCTAAAGCTCCC 1020

1021 QY AGAGCATCTCAAAAGCTGTGGTCAATGGAAACATTTATGTGGTGTGGAGGATATATG 1080  
1021 Db AGAGCATCTCAAAAGCTGTGGTCAATGGAAACATTTATGTGGTGTGGAGGATATATG 1080  
1081 QY TTCAACCACTCAGATTTAAACATGGTCTTAGCGTATGACCTTGTCTTAGGGAGTGGCTT 1140  
1081 Db TTCAACCACTCAGATTTAAACATGGTCTTAGCGTATGACCTTGTCTTAGGGAGTGGCTT 1140  
1141 QY CCACATAAACCGTCTGTGCAACAACTGTGTGTAGATATGTTCATTTCTTGGCATTTATAC 1200  
1141 Db CCACATAAACCGTCTGTGCAACAACTGTGTGTAGATATGTTCATTTCTTGGCATTTATAC 1200  
1201 QY AAGGATAAAATTTACATGTATGGAGAAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260  
1201 Db AAGGATAAAATTTACATGTATGGAGAAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260  
1261 QY TTGAGAGTTTTTCAATTCATATCAGTCATGGGTGTGTGTGACCCCTAAAGCAAGGAG 1320  
1261 Db TTGAGAGTTTTTCAATTCATATCAGTCATGGGTGTGTGTGACCCCTAAAGCAAGGAG 1320  
1321 QY CAGTATGCAAGTGGTGGGCACCTCTGCACACATTTGTACACTGAAGAAATGGCCGAGTGGTC 1380  
1321 Db CAGTATGCAAGTGGTGGGCACCTCTGCACACATTTGTACACTGAAGAAATGGCCGAGTGGTC 1380  
1381 QY ATGCTGGTCACTTTTGGTCACTGCGCTCTCTATGATATATTAAGCAATGTGACGAATAT 1440  
1381 Db ATGCTGGTCACTTTTGGTCACTGCGCTCTCTATGATATATTAAGCAATGTGACGAATAT 1440  
1441 QY GATTTGGATTAAGAACACATGAGTATATTACACACCCAGGGTGCCTTTGTGCAAGGGGT 1500  
1441 Db GATTTGGATTAAGAACACATGAGTATATTACACACCCAGGGTGCCTTTGTGCAAGGGGT 1500  
1501 QY TACGGCCATPAGCAGTGTTTACGACCATAGGACAGGGCCCTATAGCTTCATGTGGTGGTAC 1560  
1501 Db TACGGCCATPAGCAGTGTTTACGACCATAGGACAGGGCCCTATAGCTTCATGTGGTGGTAC 1560  
1561 QY AAGGCTTTTCAAGTCCAAATAGTACCGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1620  
1561 Db AAGGCTTTTCAAGTCCAAATAGTACCGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1620  
1621 QY ACCGAGTGTGACCACTTCTTAAGACAGCCGATTTTCCGTTACTTGCACACAGCTGTG 1680  
1621 Db ACCGAGTGTGACCACTTCTTAAGACAGCCGATTTTCCGTTACTTGCACACAGCTGTG 1680  
1681 QY ATAGTGAAGTGAACCATGTGGTGTGGGGGAAACACACAAATGACACATCTATGAGC 1740  
1681 Db ATAGTGAAGTGAACCATGTGGTGTGGGGGAAACACACAAATGACACATCTATGAGC 1740  
1741 QY CATGGCCCAATGTCTTCTTCCAGATTTTCATGGCCCTATGACATTTGCCCTGTGACCGCTGG 1800  
1741 Db CATGGCCCAATGTCTTCTTCCAGATTTTCATGGCCCTATGACATTTGCCCTGTGACCGCTGG 1800  
1801 QY TCAGTGTCTTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860  
1801 Db TCAGTGTCTTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTC 1860  
1861 QY TTACACACAGCACCATGTATGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1920  
1861 Db TTACACACAGCACCATGTATGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1920  
1921 QY CTGGTATTCACCTCGGAAACAGTGTGATGCGCATTCGGAGTGAAGCCGCTTTGTTAGCAGCA 1980  
1921 Db CTGGTATTCACCTCGGAAACAGTGTGATGCGCATTCGGAGTGAAGCCGCTTTGTTAGCAGCA 1980  
1981 QY GGAACCTGGTATTCGGTGTGTGGAAACAGGGTCTCTCAGTGTATCTCTGGGGGCTG 2040  
1981 Db GGAACCTGGTATTCGGTGTGTGGAAACAGGGTCTCTCAGTGTATCTCTGGGGGCTG 2040  
2041 QY GCAACTGATCAACAGAGAAAAGTTAAATCAGAATGTTTTTCCAAAAGAACTCTTTGAC 2100  
2041 Db GCAACTGATCAACAGAGAAAAGTTAAATCAGAATGTTTTTCCAAAAGAACTCTTTGAC 2100  
2101 QY CATGACAGATGTGACACAGCAACAGATTTGTTACAGCTGTACGCCAACCAATGACTGC 2160

2101 Db CATGACAGATGTGACACAGCAACAGATTTGTTACAGCTGTACAGCCAACCAATGACTGC 2160  
2161 QY CACTGTGTGAATCACCATTGTGTCCCGAGAACACACAGCTGTCTCAGAGGCCAGATCTCC 2220  
2161 Db CACTGTGTGAATCACCATTGTGTGTCCCGAGAACACACAGCTGTCTCAGAGGCCAGATCTCC 2220  
2221 QY ATTTTAGGTATGAGAAATTTGCCCCAAGGATAACCCCATGTACTACTGTAAAGAGACC 2280  
2221 Db ATTTTAGGTATGAGAAATTTGCCCCAAGGATAACCCCATGTACTACTGTAAAGAGACC 2280  
2281 QY AGCTGAGGAGCTGTCCCTGACAGACTGCGAGTGGAGGCCCGGAAATCAGAGTGC 2340  
2281 Db AGCTGAGGAGCTGTCCCTGACAGACTGCGAGTGGAGGCCCGGAAATCAGAGTGC 2340  
2341 QY ATTGCCCCGCCCCAAAATATCTGTGGCATTTGGTGGCAATTTGTTGGAATCATGTTTG 2400  
2341 Db ATTGCCCCGCCCCAAAATATCTGTGGCATTTGGTGGCAATTTGTTGGAATCATGTTTG 2400  
2401 QY AAAATTTACTACTGCAAGGAAATTTATGCAAACTGCTAAATTTGTTCTGTAGGAACCAAT 2460  
2401 Db AAAATTTACTACTGCAAGGAAATTTATGCAAACTGCTAAATTTGTTCTGTAGGAACCAAT 2460  
2461 QY GCCCTTTGGCTTCTCTTCAACCCAGAGAGTGAATTTGCTTAAAGCAGCTGCGA 2520  
2461 Db GCCCTTTGGCTTCTCTTCAACCCAGAGAGTGAATTTGCTTAAAGCAGCTGCGA 2520  
2521 QY ATATGCACTCATCTCAGACATGTCCAGCTCACCTTAACCCCATGGTGGGCTTCGG 2580  
2521 Db ATATGCACTCATCTCAGACATGTCCAGCTCACCTTAACCCCATGGTGGGCTTCGG 2580  
2581 QY AAGATCAATGTCTTACTGCTGGAAGATATGCTAAATTTGTTCTGTAGGAATTTACTA 2640  
2581 Db AAGATCAATGTCTTACTGCTGGAAGATATGCTAAATTTGTTCTGTAGGAATTTACTA 2640  
2641 QY CAGTGGATGCGCTGTAGGCCAGTGTGCTGGAATTTGTTGGAATTTTATCAGAACCCAGT 2700  
2641 Db CAGTGGATGCGCTGTAGGCCAGTGTGCTGGAATTTGTTGGAATTTTATCAGAACCCAGT 2700  
2701 QY ACTGGGAGCTGAAGGCTGCAACCTGTCATCAACCCACTCAATGCTAGTGTCTGGAAGG 2760  
2701 Db ACTGGGAGCTGAAGGCTGCAACCTGTCATCAACCCACTCAATGCTAGTGTCTGGAAGG 2760  
2761 QY CCTCAAAACACAGTGTCTAAGCAGTGTGCGGACACCATGTGCTTGGAGCAGCATGTGGA 2820  
2761 Db CCTCAAAACACAGTGTCTAAGCAGTGTGCGGACACCATGTGCTTGGAGCAGCATGTGGA 2820  
2821 QY GATTGCAACCGCGCAGCTCTGAGTGCATGTGTGAGCAACATGAAGCAGTGTGTGGAC 2880  
2821 Db GATTGCAACCGCGCAGCTCTGAGTGCATGTGTGAGCAACATGAAGCAGTGTGTGTGGAC 2880  
2881 QY TCCATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGCTATACCATGAGC 2940  
2881 Db TCCATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGCTATACCATGAGC 2940  
2941 QY ACCTGCCCCCTGAAAAATTTGTTCAAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCA 3000  
2941 Db ACCTGCCCCCTGAAAAATTTGTTCAAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCA 3000  
3001 QY GGCTGTGGTGTGTACTGATCCAGCAATCTGGCAAGGAAATGCAATAGAGGGTTC 3060  
3001 Db GGCTGTGGTGTGTACTGATCCAGCAATCTGGCAAGGAAATGCAATAGAGGGTTC 3060  
3061 QY TATAAGGACACAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATPCCACAGCCC 3120  
3061 Db TATAAGGACACAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATPCCACAGCCC 3120  
3121 QY CTGCTCAATTTCCAGCATGTGTCTAGAGACAGCAGATACACTGGTCTTTTCACTGT 3180  
3121 Db CTGCTCAATTTCCAGCATGTGTCTAGAGACAGCAGATACACTGGTCTTTTCACTGT 3180  
3181 QY CCAGCTTCCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGT 3240

Db 3181 CCAGCTTGCCAAATGCAAGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGT 3240  
QY 3241 GAGAACTGACCAAGCAGCAAGCACTGGAGACCTGATATCTGCTTACGGTATGCC 3300  
Db 3241 GAGAACTGACCAAGCAGCAAGCACTGGAGACCTGATATCTGCTTACGGTATGCC 3300  
QY 3301 ACCAATGGAGGAAATGTGAGCAGCATCAAGTGAATGGCAGCGTCTGTGCAACACC 3360  
Db 3301 ACCAATGGAGGAAATGTGAGCAGCATCAAGTGAATGGCAGCGTCTGTGCAACACC 3360  
QY 3361 AACACGGGCAAGTGTCTTGACCAACCAAGGGGCTCAAGGGGCAAGTGGCAGCTATGT 3420  
Db 3361 AACACGGGCAAGTGTCTTGACCAACCAAGGGGCTCAAGGGGCAAGTGGCAGCTATGT 3420  
QY 3421 GAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATTATCTCTTCT 3480  
Db 3421 GAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATTATCTCTTCT 3480  
QY 3481 ATTGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATC 3540  
Db 3481 ATTGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATC 3540  
QY 3541 AATTTTGTGGCTACTCTGACGACCAACCAAGGATTTGGACATGTTATCATGATGCTTCC 3600  
Db 3541 AATTTTGTGGCTACTCTGACGACCAACCAAGGATTTGGACATGTTATCATGATGCTTCC 3600  
QY 3601 AAGAATTTCAACCTCAACATCACTCGGCTGGCTGCTTCTCAGCTGGAACCCAGGCTGGA 3660  
Db 3601 AAGAATTTCAACCTCAACATCACTCGGCTGGCTGCTTCTCAGCTGGAACCCAGGCTGGA 3660  
QY 3661 GAAGATGCTGCTGTTTTCACCAACCAACATTAAGAGTACAAAGATGTTCTTCTTAAT 3720  
Db 3661 GAAGATGCTGCTGTTTTCACCAACCAACATTAAGAGTACAAAGATGTTCTTCTTAAT 3720  
QY 3721 GAGAAGTTTGAATTTGCGCAACCAACCAATATATCATCTTTTGTGTTATGTCAGTAATTC 3780  
Db 3721 GAGAAGTTTGAATTTGCGCAACCAACCAATATATCATCTTTTGTGTTATGTCAGTAATTC 3780  
QY 3781 ACTGCGCCATCAAAATTCAGGT 3803  
Db 3781 ACTGCGCCATCAAAATTCAGAT 3803

## RESULT 4

AAZ91921  
ID AAZ91921 standard; cDNA; 4072 BP.

XX AC AAZ91921;

XX AC AAZ91921;

DT 08-JUN-2000 (first entry)

DE Human mahogany protein coding sequence #3.

XX Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;  
KW weight regulation; cell therapy; body weight disorder; cachexia;  
KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.

XX Homo sapiens.

XX WO200005373-A2.

XX 03-FEB-2000.

XX 21-JUL-1999; 99WO-US016484.

XX 21-JUL-1998; 98US-0093630P.

PR 20-OCT-1998; 98US-0104978P.

PR 05-FEB-1999; 99US-00245041.

PA (MILL-) MILLENIUM PHARM INC.

PI Moore K, Nagle DL;

XX WPI; 2000-195103/17.  
DR P-PSDB; AAY81808.  
XX New human and murine mahogany genes, useful, e.g. for diagnosis and  
PT treatment of body weight disorders.  
XX Claim 1; Fig 19a; 198pp; English.  
XX This sequence represents a human mahogany gene of the invention. The  
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming  
CC therapeutics; (iii) as a source of diagnostic probes and primers for  
CC detecting expression of mg genes or mutations, regulatory defects in  
CC this gene, or for isolation of related sequences; and (iv) in (cell-  
CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to  
CC identify other (extracellular) products involved in weight regulation,  
CC and to screen for agents that disrupt interaction between (ii) and other  
CC macromolecules. The Ab are used to detect abnormal levels (or function)  
CC of (ii) (for diagnosis, prognosis or monitoring of treatment); to  
CC evaluate (ii)-expressing cells intended for cell therapy, and as  
CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
CC mg polypeptide) are used to identify agents (A) that modulate mg  
CC activity. (A) are potentially useful for the treatment of body weight  
CC disorders, particularly obesity, cachexia or anorexia, or other  
CC conditions associated with the mg gene such as hyperpigmentation,  
CC hyperphagia and disorders that result in increased metabolic rate  
XX  
SQ Sequence 4072 BP; 1087 A; 940 C; 1045 G; 1000 T; 0 U; 0 Other;

Query Match 97.2%; Score 3713.8; DB 3; Length 4072;

Best Local Similarity 98.5%; Pred. No. 0;

Matches 3761; Conservative 0; Mismatches 47; Indels 9; Gaps 1;

QY 3 GGTGGCGCAGCGCGCGCACTGAGCAGAGGCTGAGGAGGAGGAGCGCGCGCAGCAGC 62  
Db 9 GGTGGCGCAGCGCGCGCACTGAGCAGAGGCTGAGGAGGAGGAGCGCGCGCAGCAGC 68  
QY 63 GCTCGCGCGCAGGAGCGCGCGCGCGCACTGGGACGTGGGACGTGACCGAGCGCTGGAGGCG 122  
Db 69 GCTCGCGCGCAGGAGCG 128  
QY 123 GGGGCTGGGGCG 182  
Db 129 GGGACCG 188  
QY 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242  
Db 189 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 239  
QY 243 CGAGGCG 302  
Db 240 CGAGGCG 299  
QY 303 TGACCG 362  
Db 300 TGACCG 359  
QY 363 CG 422  
Db 360 CG 419  
QY 423 TTCTGGGTTTGTGACAGATGGACCTGGAAATTTATTAATACAAACGAGTGCACGTGGCT 482  
Db 420 TTCTGGGTTTGTGACAGATGGACCTGGAAATTTATTAATACAAACGAGTGCACGTGGCT 479  
QY 483 CATTGAAGGACAGCCAAATAGAAATATGAGACTTCGTTTCAATCATTTTGTACAGAGTG 542  
Db 480 CATTGAAGGACAGCCAAATAGAAATATGAGACTTCGTTTCAATCATTTTGTACAGAGTG 539  
QY 543 TAGTTGGGACCATTTATATGTTTATGATGGGAGCTCAATTTATGACCGCTAGTGTCTGC 602  
Db 540 TAGTTGGGACCATTTATATGTTTATGATGGGAGCTCAATTTATGACCGCTAGTGTCTGC 599

QY 603 ATTTAGTGGCTCATTTGTTCTGTAGAGAGATGGCAATGAGATGTCCTCCTCAGGTGTTGTC 662  
Db 600 ATTTAGTGGCTCATTTGTTCTGTAGAGAGATGGCAATGAGATGTCCTCCTCAGGTGTTGTC 659  
QY 663 CACATCAGGTATGCTTGTGCAATTTTGTAGTATGCTGCTATATAATTTGACTGATT 722  
Db 660 CACATCAGGTATGCTTGTGCAATTTTGTAGTATGCTGCTATATAATTTGACTGATT 719  
QY 723 TAATATTACTTTACAGTTTGTATATGTTCCAAATAAAGTCTCAGGCCGAGAGAGTGTAA 782  
Db 720 TAATATTACTTTACAGTTTGTATATGTTCCAAATAAAGTCTCAGGCCGAGAGAGTGTAA 779  
QY 783 GATCAGTAATAGCAGCGAACTGTTGTAATGTAATGTTCTGAAACTGGAAGGTGAAC 842  
Db 780 GATCAGTAATAGCAGCGAATCTGTTGTAATGTAATGTTCTGAAACTGGAAGGTGAAC 839  
QY 843 ATGTGACATTCCTCACTACAGCAACTGTGGTTTTCTCATCGAGCATCTCCAATTC 902  
Db 840 ATGTGACATTCCTCACTACAGCAACTGTGGTTTTCTCATCGAGCATCTCCAATTC 899  
QY 903 AAGTGATGTAGAGGATCCTCCTGTTCTCAGACTGGCAGGGTCTGTGATGTTGAGTTC 962  
Db 900 AAGTGATGTAGAGGATCCTCCTGTTCTCAGACTGGCAGGGTCTGTGATGTTGAGTTC 959  
QY 963 TGTACCACTAACCACTCATTTTGGACTCGAGAGGATATTCTAACTTAAAGCTCCCGAG 1022  
Db 960 TGTACCACTAACCACTCATTTTGGACTCGAGAGGATATTCTAACTTAAAGCTCCCGAG 1019  
QY 1023 AGCATCTCATAAAGCTGTGGTCAATGAGAAACATTATGTTGGTGTGTGAGGATATATGTT 1082  
Db 1020 AGCATCTCATAAAGCTGTGGTCAATGAGAAACATTATGTTGGTGTGTGAGGATATATGTT 1079  
QY 1083 CAAACACTCAGATTAATACATGTTCTAGCGTATGACCTTGTCTAGGGAGTGGCTTCC 1142  
Db 1080 CAAACACTCAGATTAATACATGTTCTAGCGTATGACCTTGTCTAGGGAGTGGCTTCC 1139  
QY 1143 ACTAAACCGTCTCTGTGAAACATGTTGTTAGATATGTTCTTGGCATTTATACAA 1202  
Db 1140 ACTAAACCGTCTCTGTGAAACATGTTGTTAGATATGTTCTTGGCATTTATACAA 1199  
QY 1203 GGATAAATTTACATGTTAGAGGAAATTTGATCCAACTGGGAATGTGACCAATGATT 1262  
Db 1200 GGATAAATTTACATGTTAGAGGAAATTTGATCCAACTGGGAATGTGACCAATGATT 1259  
QY 1263 GAGATTTTTCATTTCAATTAAGTATGTTGTTGAGTATGTTGAGTATGTTGAGTATGTT 1322  
Db 1260 GAGATTTTTCATTTCAATTAAGTATGTTGTTGAGTATGTTGAGTATGTTGAGTATGTT 1319  
QY 1323 GTATGAGTGGTGGGCACTCTGCACACATTTGTTACCTGAAGAATGGCCGAGTGGTCA 1382  
Db 1320 GTATGAGTGGTGGGCACTCTGCACACATTTGTTACCTGAAGAATGGCCGAGTGGTCA 1379  
QY 1383 GCTGGTCACTTTTGGTCACTGGCTCTCTATGGATATATAAGCAATGTGAGGATATGA 1442  
Db 1380 GCTGGTCACTTTTGGTCACTGGCTCTCTATGGATATATAAGCAATGTGAGGATATGA 1439  
QY 1443 TTTGGATAAGAACACATGAGTATATTAACACCCAGGGTGGCTTGTGCAAGGGGGTTA 1502  
Db 1440 TTTGGATAAGAACACATGAGTATATTAACACCCAGGGTGGCTTGTGCAAGGGGGTTA 1499  
QY 1503 CGGCCATAGCAGTGTGTTAGACCATAGGACCGAGGCCCTATAGTTTATGTTGGTACAA 1562  
Db 1500 CGGCCATAGCAGTGTGTTAGACCATAGGACCGAGGCCCTATAGTTTATGTTGGTACAA 1559  
QY 1563 GCTTTTTCAGTGCATTAAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGGATAC 1622  
Db 1560 GCTTTTTCAGTGCATTAAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGGATAC 1619  
QY 1623 CCAGATGTGACCATTTCTTAAGGACAGCGGATTTTTCCGTTACTTGCACACAGCTGTGAT 1682  
Db 1620 CCAGATGTGACCATTTCTTAAGGACAGCGGATTTTTCCGTTACTTGCACACAGCTGTGAT 1679

QY 1683 AGTGAGTGAAACCATGCTGTTGTTGGGGAACACACACATGACACATCTATGAGCA 1742  
Db 1680 AGTGAGTGAAACCATGCTGTTGTTGGGGAACACACACATGACACATCTATGAGCA 1739  
QY 1743 TGGCGCAAAATGCTTCTCTTCAGATTTTATGSCCTATGACATTTGCTGTGACCGTGTGTC 1802  
Db 1740 TGGCGCAAAATGCTTCTCTTCAGATTTTATGSCCTATGACATTTGCTGTGACCGTGTGTC 1799  
QY 1803 AGTGCTTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTTACAGAGTCTT 1862  
Db 1800 AGTGCTTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTTACAGAGTCTT 1859  
QY 1863 ACACAAACAGCAGCATGATGTGTTGGTGGTTTCAATAGTCTCTCCTCAGGCAATCT 1922  
Db 1860 ACACAAACAGCAGCATGATGTGTTGGTGGTTTCAATAGTCTCTCCTCAGGCAATCT 1919  
QY 1923 GGTATTCACTCGGAAACAGTGTGATCGCATCGAGTGAAGCGCTGTTGTTAGCAGCAG 1982  
Db 1920 GGTATTCACTCGGAAACAGTGTGATCGCATCGAGTGAAGCGCTGTTGTTAGCAGCAG 1979  
QY 1983 ACCTGCTATTTCGTTGTGTGGAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGTC 2042  
Db 1980 ACCTGCTATTTCGTTGTGTGGAACACAGGGTCTCTCAGTGTATCTCGTGGCGCTGTC 2039  
QY 2043 AACTGATGAACAGAAAGAAAGTTAAATCAGATGTTTTCAAAAGAACTCTTGACCA 2102  
Db 2040 AACTGATGAACAGAAAGAAAGTTAAATCAGATGTTTTCAAAAGAACTCTTGACCA 2099  
QY 2103 TGACAGATGTGACACACACAGATTTTACAGCTGTACAGCCACACACAACTGACTGCCA 2162  
Db 2100 TGACAGATGTGACACACACAGATTTTACAGCTGTACAGCTGCACAGCCACACAACTGACTGCCA 2159  
QY 2163 CTGCTGCAATGACCATTTGTTCTCCAGGAAACACAGCTGCTCAGAGGCGCAGATCTCCAT 2222  
Db 2160 CTGCTGCAATGACCATTTGTTCTCCAGGAAACACAGCTGCTCAGAGGCGCAGATCTCCAT 2219  
QY 2223 TTTTAGTATGAGAAATGCCCCAAGATTAACCCATGTAATCTGTAAACAGAGAACCCAG 2282  
Db 2220 TTTTAGTATGAGAAATGCCCCAAGATTAACCCATGTAATCTGTAAACAGAGAACCCAG 2279  
QY 2283 CTGACAGAGTGTGCTGACACAGAACTGCCAGTGGGAGCCCGGAATCAGGAGTGCAT 2342  
Db 2280 CTGACAGAGTGTGCTGACACAGAACTGCCAGTGGGAGCCCGGAATCAGGAGTGCAT 2339  
QY 2343 TGGCTTCCCGGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACCTCATGTTGAA 2402  
Db 2340 TGGCTTCCCGGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACCTCATGTTGAA 2399  
QY 2403 AATTACTACTGCCAAGGAGAAATATGACAAATGCTAAATCTTCTGTAGGAACCAACAATGC 2462  
Db 2400 AATTACTACTGCCAAGGAGAAATATGACAAATGCTAAATCTTCTGTAGGAACCAACAATGC 2459  
QY 2463 CCTTTTGGCTTCTCTTACACCCAGAAAGGTAGAAATTTGCTTAAAGCAGCTGCGAAT 2522  
Db 2460 CCTTTTGGCTTCTCTTACACCCAGAAAGGTAGAAATTTGCTTAAAGCAGCTGCGAAT 2519  
QY 2523 AATCAGTATCTCAGAGCATGTCAGCTCACTTAACCCCATGGTCCGCTTCGGAA 2582  
Db 2520 AATCAGTATCTCAGAGCATGTCAGCTCACTTAACCCCATGGTCCGCTTCGGAA 2579  
QY 2583 GATCAATGTGCTTACTGTTGCTGGGAAGATATGTCCTCCATTTTACAAATAGTTTACTACA 2642  
Db 2580 GATCAATGTGCTTACTGTTGCTGGGAAGATATGTCCTCCATTTTACAAATAGTTTACTACA 2639  
QY 2643 GTGATGCTGCTGAGCCAGTGTGCTGGATTTCTGTGGAAATTTTATCAGAACCCAGTAC 2702  
Db 2640 GTGATGCTGCTGAGCCAGTGTGCTGGATTTCTGTGGAAATTTTATCAGAACCCAGTAC 2699  
QY 2703 TCGGGAGTGAAGCTGCAACCTGCACTCAACCCACTCAATGGTGTGTTCTGTGAAAGGCC 2762  
Db 2700 TCGGGAGTGAAGCTGCAACCTGCACTCAACCCACTCAATGGTGTGTTCTGTGAAAGGCC 2759  
QY 2763 TGCAAAACAGTGTGAAGCAGTCCCGACACCACTATGCTCCTTGAAGCAGCATGTGAGAA 2822









QY 2343 TGCCCTGCCGAAAAATATCTGTGGCATTTGGCTGGCAATTTGGTGGAAACTCATGTTTGA 2402  
Db 2340 TGCCCTGCCGAAAAATATCTGTGGCATTTGGCTGGCAATTTGGTGGAAACTCATGTTTGA 2399  
QY 2403 AATTACTACTGCCAAGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGAACCAACAATGC 2462  
Db 2400 AATTACTACTGCCAAGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGAACCAACAATGC 2459  
QY 2463 CTTTTGGCTTCTCTTACACCCAGAGAGAGTAAATTTGCTTTAAAGCAGCTGGAT 2522  
Db 2460 CTTTTGGCTTCTCTTACACCCAGAGAGAGTAAATTTGCTTTAAAGCAGCTGGAT 2519  
QY 2523 AATGCACTCATCTCAGAGCATGTCCTCAAGCTCACCTTTAACCCTATGGCTGGCTTCGGAA 2582  
Db 2520 AATGCACTCATCTCAGAGCATGTCCTCAAGCTCACCTTTAACCCTATGGCTGGCTTCGGAA 2579  
QY 2583 GATCAATGTCTCTACTGTCTGGGAAGATATGTCCTCAATTTACAAATAGTTTACTACA 2642  
Db 2580 GATCAATGTCTCTACTGTCTGGGAAGATATGTCCTCAATTTACAAATAGTTTACTACA 2639  
QY 2643 GTGGATGCCCTCTGAGCCAGTCAATGCTGGAATTTCTGGAATTTTATCAGAACCCAGTAC 2702  
Db 2640 GTGGATGCCCTCTGAGCCAGTCAATGCTGGAATTTCTGGAATTTTATCAGAACCCAGTAC 2699  
QY 2703 TCGGGGACTGGAAGCTGCAACTGTGCATCAACCCACTCAATGTGTGCTGTGAAAGGCC 2762  
Db 2700 TCGGGGACTGGAAGCTGCAACTGTGCATCAACCCACTCAATGTGTGCTGTGAAAGGCC 2759  
QY 2763 TGCARACCAAGTCTAAGCAGTGGCGGACACCATGTGCTTGGAGCAGCATGTGGAGA 2822  
Db 2760 TGCARACCAAGTCTAAGCAGTGGCGGACACCATGTGCTTGGAGCAGCATGTGGAGA 2819  
QY 2823 TTGCACAGCGGAGCTCTGAGTGCATGTGTGAGCAACATGAAGCAGTGTGTGCACTC 2882  
Db 2820 TTGCACAGCGGAGCTCTGAGTGCATGTGTGAGCAACATGAAGCAGTGTGTGCACTC 2879  
QY 2883 CAATGCTATGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTATACATGAGCAGC 2942  
Db 2880 CAATGCTATGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTATACATGAGCAGC 2939  
QY 2943 CTGCCCCCTGAAAATTTGTCAGGCTACTGTACTGTACTGTACTGTACTGTACTGTACTGT 3002  
Db 2940 CTGCCCCCTGAAAATTTGTCAGGCTACTGTACTGTACTGTACTGTACTGTACTGTACTGT 2999  
QY 3003 CTGTGCTGGTGTACTGATCCAGCAATCTGTGGCAAGAGGAATGCATGAGGGTTCCTA 3062  
Db 3000 CTGTGCTGGTGTACTGATCCAGCAATCTGTGGCAAGAGGAATGCATGAGGGTTCCTA 3059  
QY 3063 TAAAGGACCAAGTGAAGATGCTTGGCAAGCCCTACAGAAATTTCTATCCACAGCCCTC 3122  
Db 3060 TAAAGGACCAAGTGAAGATGCTTGGCAAGCCCTACAGAAATTTCTATCCACAGCCCTC 3119  
QY 3123 GTCATATTCAGCATGTGTCTAGAGCAGCAGATACAACTGGTCTTTCAITTCATGTC 3182  
Db 3120 GTCATATTCAGCATGTGTCTAGAGCAGCAGATACAACTGGTCTTTCAITTCATGTC 3179  
QY 3183 AGCTTGCCAAATCAAGGCCAGTAAATGATCAATCAGAGCATCTGTGAGAGTGTGA 3242  
Db 3180 AGCTTGCCAAATCAAGGCCAGTAAATGATCAATCAGAGCATCTGTGAGAGTGTGA 3239  
QY 3243 GAACCTGACCAAGCAAGCACTGCGAGACCTGCAATCTGGCTTCTACGGTGTATCCAC 3302  
Db 3240 GAACCTGACCAAGCAAGCACTGCGAGACCTGCAATCTGGCTTCTACGGTGTATCCAC 3299  
QY 3303 CAATGGAGGAATCTCAGCCATGCAAGTGCATGGAATGGGACCGCTCTGTGCAACACCAA 3362  
Db 3300 CAATGGAGGAATCTCAGCCATGCAAGTGCATGGAATGGGACCGCTCTGTGCAACACCAA 3359  
QY 3363 CACGGGCAAGTCTTCTGCAACCAAGGGGCTCAAGGGGACGAGTCCAGCTATGTGA 3422  
Db 3360 CACGGGCAAGTCTTCTGCAACCAAGGGGCTCAAGGGGACGAGTCCAGCTATGTGA 3419  
QY 3423 GGTAGAAAATGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATCTCTTTAT 3482

Db 3420 GGTAGAAAATGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATCTCTTTAT 3479  
QY 3483 TGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAA 3542  
Db 3480 TGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAA 3539  
QY 3543 TTTTGTGGCTACTCTCAGCAACCAACAGGATTTGGACATGTTTCATCAATGCCTCCAA 3602  
Db 3540 TTTTGTGGCTACTCTCAGCAACCAACAGGATTTGGACATGTTTCATCAATGCCTCCAA 3599  
QY 3603 GAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3662  
Db 3600 GAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3659  
QY 3663 AGAGATGCTGTGTTTTCACAAACCAACATTAAGGAGTACAAAGATGTTTCTTAATGA 3722  
Db 3660 AGAGATGCTGTGTTTTCACAAACCAACATTAAGGAGTACAAAGATGTTTCTTAATGA 3719  
QY 3723 GAAGTTTGAATTTTCGCAACCAACCAATATCACCTTTCTTTGTTTATGTCAGTAATTTCC 3782  
Db 3720 GAAGTTTGAATTTTCGCAACCAACCAATATCACCTTTCTTTGTTTATGTCAGTAATTTCC 3779  
QY 3783 CTGCCCCATCAAAATTCAGT 3803  
Db 3780 CTGCCCCATCAAAATTCAGAT 3800

RESULT 6  
AAZ52275  
ID AAZ52275 standard; cDNA; 4068 BP.  
XX  
AC AAZ52275;  
XX  
DT 18-JUL-2000 (first entry)  
XX  
DE Human membrane attractin-1 cDNA.  
XX  
KW Human; membrane attractin-1; immune response; macrophage; monocyte;  
KW T cell; immunostimulant; immunosuppressed patient; cancer;  
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..4068  
FT /tag= a  
FT /product= "Membrane attractin-1"  
XX  
FN W0200015651-A1.  
XX  
PD 23-MAR-2000.  
XX  
PF 14-SEP-1999; 99WO-US020948.  
XX  
PR 14-SEP-1998; 98US-0100137P.  
XX  
PA (DAND ) DANA FARBER CANCER INST INC.  
XX  
PI Duke-Cohan JS, Schlossman SF;  
DR WPI; 2000-271373/23.  
DR P-PSDB; AAY70690.  
XX  
PT Isolated nucleic acids encoding human attractin polypeptides useful for  
PT enhancing immune responses.  
XX  
PS Claim 3; Fig 10; 120pp; English.  
XX  
CC The patent discloses four forms of human attractin polypeptides which  
CC enhance immune response by promoting macrophage and monocyte spreading in  
CC the presence of T cells. These include soluble attractin-1 and -2 and  
CC membrane attractin-1 and -2. These various forms of attractin are encoded

CC by alternatively spliced mRNA molecule transcribed from a single gene.  
CC The present sequence is a cDNA encoding human membrane attractin-1.  
CC Membrane attractin differs from soluble attractin in having a  
CC transmembrane domain and a cytoplasmic domain at the C-terminal. This  
CC sequence can be used to enhance immune response in immunosuppressed  
CC patients such as those undergoing chemo- and radio-therapy treatment for  
CC cancer or those suffering from common variable immunodeficiency syndrome.  
CC The proteins may also be used to screen modulators (agonists and  
CC antagonists) of immune response which may also be used to regulate immune  
CC reactions. Attractin antibodies can be used to inhibit immune response in  
CC transplant recipients or patients afflicted with autoimmune disease  
XX  
SQ Sequence 4068 BP; 1042 A; 948 C; 1027 G; 1051 T; 0 U; 0 Other;  
Query Match 91.4%; Score 3492.4; DB 3; Length 4068;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 3508; Conservative 0; Mismatches 26; Indels 0; Gaps 0;  
Qy 270 GTCGGGCTCAGCCGACGCGGCTCGCGGCGAGGCGCGCGGCTGTGTCAACGGCGGTG 329  
Db 48 GCGCGGACGCGGCGCTCGCGGCGAGGCGCGCGGCGGCTGTGTCAACGGCGGTG 107  
Qy 330 CTGCAACCCCTGGCACCGCGGCGAGTGCCTGCGCGCGCGGCTGGTGGCGGAGCAATGCCA 389  
Db 108 CTGCAACCCCTGGCACCGCGGCGAGTGCCTGCGCGCGCGGCTGGTGGCGGAGCAATGCCA 167  
Qy 390 GCACTGGGGGGCGCGCTTCAGACTAACTGGATCTTCTGGGTTGTGACAGATGACCTGG 449  
Db 168 GCACTGGGGGGCGCGCTTCAGACTAACTGGATCTTCTGGGTTGTGACAGATGACCTGG 227  
Qy 450 AAATTTAAATACAAAAGAGTGCAGTGGCTCATTTGAGGACAGCCAAATGAATAAT 509  
Db 228 AAATTTAAATACAAAAGAGTGCAGTGGCTCATTTGAGGACAGCCAAATGAATAAT 287  
Qy 510 GAGACTTCGTTTCAATCAATTTGTACAGAGTGTAGTTGGACCATTTATATGTTTATGA 569  
Db 288 GAGACTTCGTTTCAATCAATTTGTACAGAGTGTAGTTGGACCATTTATATGTTTATGA 347  
Qy 570 TGGGACTCAATTTATGACCGCTAGTTGTGCGAATTTAGTGGCCCTCATTTGCTCTGAGAG 629  
Db 348 TGGGACTCAATTTATGACCGCTAGTTGTGCGAATTTAGTGGCCCTCATTTGCTCTGAGAG 407  
Qy 630 AGATGGCAATGAGACTGTCCTGAGTTGTTGCCACATCAGTTTATGCTTGTGCAATTT 689  
Db 408 AGATGGCAATGAGACTGTCCTGAGTTGTTGCCACATCAGTTTATGCTTGTGCAATTT 467  
Qy 690 TTTTATGATGCTGCTTAAATTTGACTCGAATTTAATAATTAATTAATTAATTAATGATG 749  
Db 468 TTTTATGATGCTGCTTAAATTTGACTCGAATTTAATAATTAATTAATTAATTAATGATG 527  
Qy 750 TCCAAATTAATCTCAGCGCGGAGAGTGTAGATCAGTAATAGCAGCGAACTGTTGA 809  
Db 528 TCCAAATTAATCTCAGCGCGGAGAGTGTAGATCAGTAATAGCAGCGAATCTGTTGA 587  
Qy 810 ATGTGAATTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869  
Db 588 ATGTGAATTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647  
Qy 870 CTGTGGTTTCTCATGAGGCACTGCAATTCAGATGATGACAGATGCTCTGCTT 929  
Db 648 CTGTGGTTTCTCATGAGGCACTGCAATTCAGATGATGACAGATGCTCTGCTT 707  
Qy 930 CTCAGACTGGCAGGTCCTGGATGTTTCACTTCTGTACCAGTAAACCACTCATTTTGGAC 989  
Db 708 CTCAGACTGGCAGGTCCTGGATGTTTCACTTCTGTACCAGTAAACCACTCATTTTGGAC 767  
Qy 990 TCGAGAGGAATATTTCACTTAAAGCTCCCGAGGATCTCAATAAGCTGTGGTCAATGG 1049  
Db 768 TCGAGAGGAATATTTCACTTAAAGCTCCCGAGGATCTCAATAAGCTGTGGTCAATGG 827  
Qy 1050 AAACATTAATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTTATACATGTTCT 1109  
Db 828 AAACATTAATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTTATACATGTTCT 887

Qy 1110 AGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCGTCTCTGTGAACAATGTGGT 1169  
Db 888 AGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCGTCTCTGTGAACAATGTGGT 947  
Qy 1170 TGTAGATATGGTCAATCTTTGGCATATACAAGATAAAATTTACATGTATGGAGGAAA 1229  
Db 948 TGTAGATATGGTCAATCTTTGGCATATACAAGATAAAATTTACATGTATGGAGGAAA 1007  
Qy 1230 AATTGATCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATAATGAGTC 1289  
Db 1008 AATTGATCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATAATGAGTC 1067  
Qy 1290 ATGGGTGTTTGTGACCCCTTAGGCAAGAGCAGTATGAGTGGTGGGCACTCTGCAACA 1349  
Db 1068 ATGGGTGTTTGTGACCCCTTAGGCAAGAGCAGTATGAGTGGTGGGCACTCTGCAACA 1127  
Qy 1350 CAATTGTTACACTGAAGAATGGCCGAGTGGTCAATGCTGTGTCATCTTTGGTCACTGCGCTCT 1409  
Db 1128 CAATTGTTACACTGAAGAATGGCCGAGTGGTCAATGCTGTGTCATCTTTGGTCACTGCGCTCT 1187  
Qy 1410 CTATGGATATATAAGCAATGTGACGAATATGATTTGATTAAGAACACATGAGGATATAT 1469  
Db 1188 CTATGGATATATAAGCAATGTGACGAATATGATTTGATTAAGAACACATGAGGATATAT 1247  
Qy 1470 ACACACCCAGGCTGCTTGTGCAAGGGGTTACGGCCATAGCAGTGTGTTTACGACCATAG 1529  
Db 1248 ACACACCCAGGCTGCTTGTGCAAGGGGTTACGGCCATAGCAGTGTGTTTACGACCATAG 1307  
Qy 1530 GACCAGGGCCCTATAGCTTCAATGAGTGGCTTCAAGGCTTTCAGTGCCTAATAAGTACGGCT 1589  
Db 1308 GACCAGGGCCCTATAGCTTCAATGAGTGGCTTCAAGGCTTTCAGTGCCTAATAAGTACGGCT 1367  
Qy 1590 TGCAGATGATCTCTACCGATATGATGGATACCCAGATGTGACCATTCCTTAAAGGACAG 1649  
Db 1368 TGCAGATGATCTCTACCGATATGATGGATACCCAGATGTGACCATTCCTTAAAGGACAG 1427  
Qy 1650 CCGATTTTCCGTTACTTGCACACACAGCTGTGATAGTGGAGCAACCATCTGCTGTTGG 1709  
Db 1428 CCGATTTTCCGTTACTTGCACACACAGCTGTGATAGTGGAGCAACCATCTGCTGTTGG 1487  
Qy 1710 GGAACAACACACATGACATCTATAGCCATGAGCCGCAAAATGCTTCTTTCAGATTT 1769  
Db 1488 GGAACAACACACATGACATCTATAGCCATGAGCCGCAAAATGCTTCTTTCAGATTT 1547  
Qy 1770 CATGGCTATGACATTCCTGACCGCTGGTCAAGTCTTCCAGACCTGATCTCCACCA 1829  
Db 1548 CATGGCTATGACATTCCTGACCGCTGGTCAAGTCTTCCAGACCTGATCTCCACCA 1607  
Qy 1830 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACACAGCACCATGTATGTGTCGG 1889  
Db 1608 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACACAGCACCATGTATGTGTCGG 1667  
Qy 1890 TGGTTTCAATAGTCTCTCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGC 1949  
Db 1668 TGGTTTCAATAGTCTCTCTCAGCGACATCCTGGTATTCACCTCGGAACAGTGTGATGC 1727  
Qy 1950 GCATCGAGTGAAGCGCGCTTGTGTTAGCAGCAGGACCTGGTATTCGGTGTGTGGAACAC 2009  
Db 1728 GCATCGAGTGAAGCGCGCTTGTGTTAGCAGCAGGACCTGGTATTCGGTGTGTGGAACAC 1787  
Qy 2010 AGGTGCTCTCAGTGTATCTCTGGGGCTGGCAACTGATGAACAGAAAGGTTAAA 2069  
Db 1788 AGGTGCTCTCAGTGTATCTCTGGGGCTGGCAACTGATGAACAGAAAGGTTAAA 1847  
Qy 2070 ATCAGATGTTTTTCCAAAAGAACTCTTTGACCATGACAGATGTGACGACACAGATTG 2129  
Db 1848 ATCAGATGTTTTTCCAAAAGAACTCTTTGACCATGACAGATGTGACGACACAGATTG 1907  
Qy 2130 TTACAGCTGTACAGCCAAACCAATGATGCTGCCATGCGCAATGACCATTTGTGTCCCCAG 2189  
Db 1908 TTACAGCTGTACAGCCAAACCAATGATGCTGCCATGCGCAATGACCATTTGTGTCCCCAG 1967

QY 2190 GAACACAGCTGCTCAGAGGCGCAGATCTCCATTTTATAGTATGAGATTCGCCCAAGGA 2249  
Db 1968 GAACCAACAGCTGCTCAGAGGCGCAGATCTCCATTTTATAGTATGAGATTCGCCCAAGGA 2027  
QY 2250 TAAACCCATGTACTACTGTAAACAAGAACACAGCTGCGAGGAGCTGTGCCCTGGACAGAA 2309  
Db 2028 TAAACCCATGTACTACTGTAAACAAGAACACAGCTGCGAGGAGCTGTGCCCTGGACAGAA 2087  
QY 2310 CTGCCAGTGGGAGCCCCGGAATCAGAGTGCAATTCGCCCTGCCCGAAATATCTGTGGCAT 2369  
Db 2088 CTGCCAGTGGGAGCCCCGGAATCAGAGTGCAATTCGCCCTGCCCGAAATATCTGTGGCAT 2147  
QY 2370 TGGCTGGCAATTTGGTTTGGAAATCTCATGTTTGAATTAATCTACTGCGCAAGGAGAAATTAAGA 2429  
Db 2148 TGGCTGGCAATTTGGTTTGGAAATCTCATGTTTGAATTAATCTACTGCGCAAGGAGAAATTAAGA 2207  
QY 2430 CAATGCTAAATTTGTTCTGTAGGAACCAACAATGCCCTTTGGCTTCTCTTTAAACCCAGAA 2489  
Db 2208 CAATGCTAAATTTGTTCTGTAGGAACCAACAATGCCCTTTGGCTTCTCTTTAAACCCAGAA 2267  
QY 2490 GAAGGTAGAAATTTGCTTAAAGCAGCTGCGAATATGCAATCTCAGTCACTCAGAGCATGTCCAA 2549  
Db 2268 GAAGGTAGAAATTTGCTTAAAGCAGCTGCGAATATGCAATCTCAGTCACTCAGAGCATGTCCAA 2327  
QY 2550 GGTCAACCTTAACCCCATGGTGGGCTTTGGAAGATCAATGTGCTCTACTGGTCTGGGA 2609  
Db 2328 GGTCAACCTTAACCCCATGGTGGGCTTTGGAAGATCAATGTGCTCTACTGGTCTGGGA 2387  
QY 2610 AGATATGCTCCCATTTACAAATAGTTTACTACAGTGGATGCGTCTGAGCCAGTATGC 2669  
Db 2388 AGATATGCTCCCATTTACAAATAGTTTACTACAGTGGATGCGTCTGAGCCAGTATGC 2447  
QY 2670 TGGATTTCTGTGAAATTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACTGCAT 2729  
Db 2448 TGGATTTCTGTGAAATTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACTGCAT 2507  
QY 2730 CAACCCATCAATGTGTAGTGTCTGAAAGGCTTGCAAAACACAGTGTCTAAGCAGTCCGG 2789  
Db 2508 CAACCCATCAATGTGTAGTGTCTGAAAGGCTTGCAAAACACAGTGTCTAAGCAGTCCGG 2567  
QY 2790 GACACCATGTGCTTGAGACAGCATGTGGAGATTGGACAGCGCGAGCTCTGAGTGCAT 2849  
Db 2568 GACACCATGTGCTTGAGACAGCATGTGGAGATTGGACAGCGCGAGCTCTGAGTGCAT 2627  
QY 2850 GTGTGCGAGCAACATGAAGCAGTGTGTGAGTCTCAATGCCCTATGTGCCCTCTTCCCTTT 2909  
Db 2628 GTGTGCGAGCAACATGAAGCAGTGTGTGAGTCTCAATGCCCTATGTGCCCTCTTCCCTTT 2687  
QY 2910 TGGCCAGTGTATGGAATCGTATACGATGAGCACCTGCCCTGAAATTTGTCAGGCTA 2969  
Db 2698 TGGCCAGTGTATGGAATCGTATACGATGAGCACCTGCCCTGAAATTTGTCAGGCTA 2747  
QY 2970 CTGTACCTGTAGTCAATTTGAGCAACACAGGCTGTGGCTGTGTACTGATCCACAGCA 3029  
Db 2748 CTGTACCTGTAGTCAATTTGAGCAACACAGGCTGTGGCTGTGTACTGATCCACAGCA 2807  
QY 3030 TACTGGCAAGGAAATGATAGAGGTTCTCTATAAAGACCAAGTGAAGATGCTTCGCA 3089  
Db 2808 TACTGGCAAGGAAATGATAGAGGTTCTCTATAAAGACCAAGTGAAGATGCTTCGCA 2867  
QY 3090 AGCCCTTACAGGAAATTTCTATCACAGACCCCTGCTCAATTCACAGCATGTGTCTAGAGGA 3149  
Db 2868 AGCCCTTACAGGAAATTTCTATCACAGACCCCTGCTCAATTCACAGCATGTGTCTAGAGGA 2927  
QY 3150 CAGCAGATACACTGGTCTTCTATCTCACTGTCCAGTGTGCAATGCAACGGCCACAGTAA 3209  
Db 2928 CAGCAGATACACTGGTCTTCTATCTCACTGTCCAGTGTGCAATGCAACGGCCACAGTAA 2987  
QY 3210 ATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACTTCACACAGGCAAGCACTGCGA 3269  
Db 2988 ATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACTTCACACAGGCAAGCACTGCGA 3047  
QY 3270 GACCTGCATATCTGGCTTCTACGGTGTATCCCAACAAATGGAGGGAATGTCTAGCCATGCA 3329

Db 3048 GACCTGCATATCTGGCTTCTACGGTATCCCAATGGAGGAAATGTCTAGCCATGCA 3107  
QY 3330 GTGCAATGGGCAACGGCTCTCTGTGCAACACCAACAGGCAAGTGTCTTGACACCAAA 3389  
Db 3108 GTGCAATGGGCAACGGCTCTCTGTGCAACACCAACAGGCAAGTGTCTTGACACCAAA 3167  
QY 3390 GGGCGTCAAGGGGGAAGAGTGGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCC 3449  
Db 3168 GGGCGTCAAGGGGGAAGAGTGGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCC 3227  
QY 3450 TCTCAGAGGAACATGTTATTTATCTCTTATTGACTATCAGTTCACCTTTAGTCTATC 3509  
Db 3228 TCTCAGAGGAACATGTTATTTATCTCTTATTGACTATCAGTTCACCTTTAGTCTATC 3287  
QY 3510 CCAGGAAGATGATCGCTATTATACACAGCTATCAATTTTGGCTACTCTCTGAGAAACAAA 3569  
Db 3288 CCAGGAAGATGATCGCTATTATACACAGCTATCAATTTTGGCTACTCTCTGAGAAACAAA 3347  
QY 3570 CAGGGATTTGGACATCTTCATCAATGCCCTCAAGATTTCAACCTCAACATCACCTGGGC 3629  
Db 3348 CAGGGATTTGGACATCTTCATCAATGCCCTCAAGATTTCAACCTCAACATCACCTGGGC 3407  
QY 3630 TGGCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGATGCTGTGTTTCAAAAACCAA 3689  
Db 3408 TGGCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGATGCTGTGTTTCAAAAACCAA 3467  
QY 3690 CATTAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTGATTTTGGCAACCAACCCAAA 3749  
Db 3468 CATTAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTGATTTTGGCAACCAACCCAAA 3527  
QY 3750 TATCACTTTCTTTGTTTATGTAGTCAATAATTTCACTGGCCCATCAAAATTCAGST 3803  
Db 3528 TATCACTTTCTTTGTTTATGTAGTCAATAATTTCACTGGCCCATCAAAATTCAGAT 3581

## RESULT 7

AAZ52274

ID AAZ52274 standard; cDNA; 3597 Bp.

XX AAZ52274;

XX AC

XX DT -18-JUL-2000 (first entry)

XX DE Human soluble attractin-1 cDNA.

XX KW Human; soluble attractin-1; immune response; macrophage; monocyte;  
XX KW T cell; immunostimulant; immunosuppressed patient; cancer;  
XX KW immunodeficiency syndrome; transplant; autoimmune disease; ss.

XX OS Homo sapiens.

XX FH Key

XX FT Location/Qualifiers

XX FT 1. 3597

XX FT /tag= a

XX FT /product= "soluble attractin-1"

XX PN WO200015651-A1.

XX PD 23-MAR-2000.

XX PF 14-SEP-1999; 99WO-US020948.

XX PF 14-SEP-1999; 98US-0100137P.

XX PR (DAND ) DANA PARBER CANCER INST INC.

XX PA Duke-Cohan JS, Schlosman SF;

XX PI WPI; 2000-271373/23.

XX DR P-PSDB; AAY70689.

XX PT Isolated nucleic acids encoding human attractin polypeptides useful for

PT	enhancing immune responses.	990	TCGAGAGGAATATTTAACTTAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGG	1049
XX	Claim 3; Fig 8; 120pp; English.	768	TCGAGAGGAATATTTAACTTAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGG	827
PS		1050	AAACATATATGTGGTGTGGAGGATATATTTCAACCACTCAGATTATAACATGTTCT	1109
XX	The patent discloses four forms of human attractin polypeptides which	828	AAACATATATGTGGTGTGGAGGATATATTTCAACCACTCAGATTATAACATGTTCT	887
CC	enhance immune response by promoting macrophage and monocyte spreading in	1110	AGCGTATGACCTTGTCTTAGGAGTGGTTCACATAAACCGTCTGTGAAAGATGCT	1169
CC	the presence of T cells. These include soluble attractin-1 and -2 and	888	AGCGTATGACCTTGTCTTAGGAGTGGTTCACATAAACCGTCTGTGAAAGATGCT	947
CC	membrane attractin-1 and -2. These various forms of attractin are encoded	1170	TGTTAGATATGTCATCTTTGGCATTATACAAAGGATAAAATTTACATGTATGGAGAAA	1229
CC	by alternatively spliced mRNA molecule transcribed from a single gene.	948	TGTTAGATATGTCATCTTTGGCATTATACAAAGGATAAAATTTACATGTATGGAGAAA	1007
CC	The present sequence is a cDNA encoding soluble attractin-1 obtained from	1230	AATTGATCCAACTGGGAATGTGCCAATGAGTTGAGAGTTTTCACATTCATAATGAGTC	1289
CC	activated T cell and human fetal liver libraries. This sequence can be	1008	AATTGATCCAACTGGGAATGTGCCAATGAGTTGAGAGTTTTCACATTCATAATGAGTC	1067
CC	used to enhance immune response in immunosuppressed patients such as	1290	ATGGGTGTTGTGACCCCTTAAGCAAGGAGGAGTATGAGTGGTGGGCACTCTGCACA	1349
CC	those undergoing chemo- and radio-therapy treatment for cancer or those	1068	ATGGGTGTTGTGACCCCTTAAGCAAGGAGGAGTATGAGTGGTGGGCACTCTGCACA	1127
CC	suffering from common variable immunodeficiency syndrome. The proteins	1350	CATTGTTACACTGAAGAATGGCCGAGTGTCTATGCTGCTCATCTTTGGTCACTGCCCTCT	1409
CC	may also be used to screen modulators (agonists and antagonists) of	1128	CATTGTTACACTGAAGAATGGCCGAGTGTCTATGCTGCTCATCTTTGGTCACTGCCCTCT	1187
CC	immune responses which may be used to regulate immune reactions.	1410	CTATGGATATATAGCAATGTGCAGGAATATGATTTGGATAGAAACACATGGAGTATAT	1469
CC	Attractin antibodies can be used to inhibit immune response in transplant	1188	CTATGGATATATAGCAATGTGCAGGAATATGATTTGGATAGAAACACATGGAGTATAT	1247
CC	recipients or patients afflicted with autoimmune disease	1470	ACACACCCAGGGTGGCCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTTCACGACCATAG	1529
XX	Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;	1248	ACACACCCAGGGTGGCCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTTCACGACCATAG	1307
SQ	Query Match 91.3%; Score 3487; DB 3; Length 3597;	1530	GACCAGGGCCCTATACGTTTCATGTTGGCTACAGGCTTTTCACTGCCAATAGTACCGGCT	1589
	Best Local Similarity 99.2%; Pred. No. 0;	1308	GACCAGGGCCCTATACGTTTCATGTTGGCTACAGGCTTTTCACTGCCAATAGTACCGGCT	1367
	Matches 3524; Conservative 0; Mismatches 25; Indels 2; Gaps 2;	1590	TGCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGGACCACTCTTAAAGACAG	1649
QY	270 GTCCGGCTCAGCCGAGCGAGCCGCAAGGAATGTGACCGGCCCTGTGTCAACGGCGTCG	1368	TGCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGGACCACTCTTAAAGACAG	1427
DB	48 GCGCGGACCGCAGCGCTCGCGGCGAGGCGCGCGGCGCACTGTGTCAACGGCGTCG	1650	CCGATTTTTCGTTTACCTTGACACAGCTGTGTGATGTAGTGGTGGAAACCATGCTGTGTTGG	1709
QY	330 CTGCAACCTGACCGCGCAGTGTCTGCTGCCCGCGGCTGGGTGGCGGAGCAATGCCA	1428	CCGATTTTTCGTTTACCTTGACACAGCTGTGTGATGTAGTGGTGGAAACCATGCTGTGTTGG	1487
DB	108 CTGCAACCTGACCGCGCAGTGTCTGCTGCCCGCGGCTGGGTGGCGGAGCAATGCCA	1710	GGGAAACACACAAATGACATCTATGAGCCATGGCGCCAAATGCTTCTTTCAGATTT	1769
QY	390 GCATCTGCGGGCGCTTTCAGACTAACTGGATCTTCTGGTGTGTGAAGATGACCTGG	1488	GGGAAACACACAAATGACATCTATGAGCCATGGCGCCAAATGCTTCTTTCAGATTT	1547
DB	168 GCATCTGCGGGCGCTTTCAGACTAACTGGATCTTCTGGTGTGTGAAGATGACCTGG	1770	CATGCCCTATGACATTCGCTGTGACCGCTGTGACCGCTGGTGGTGGTGGTGGTGGTGG	1829
QY	450 AAATATTAATAACAACGAACTGACGCTGCTTGAAGACACGCAATAGATAAT	1548	CATGCCCTATGACATTCGCTGTGACCGCTGTGACCGCTGGTGGTGGTGGTGGTGGTGG	1606
DB	228 AAATATTAATAACAACGAACTGACGCTGCTTGAAGACACGCAATAGATAAT	1830	TGATGTCAACAGATTTGGCCATT - CAGCAGTCTTTACACACAGCAGCAGCAGTGTGTCG	1888
QY	510 GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGGACCACTTTATATGA	1607	TGATGTCAACAGATTTGGCCATT - CAGCAGTCTTTACACACAGCAGCAGCAGTGTGTCG	1666
DB	288 GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGGACCACTTTATATGA	1889	GTGGTTTCAATAGTCTCTCTCTCAGCAGCATCTCTGTTTACACCTCGGAAACAGTGTGATG	1948
QY	570 TGGGACTCAATTTATGCAACCGCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1667	GTGGTTTCAATAGTCTCTCTCTCAGCAGCATCTCTGTTTACACCTCGGAAACAGTGTGATG	1726
DB	348 TGGGACTCAATTTATGCAACCGCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1949	CGCATCGGAGTGAAGAGCGCTTTGTTAGCAGCAGGACCTGGTATTCGGTGTGTGGAAACA	2008
QY	630 AGATGGCAATGAGACTGCTCCCTGAGGTGTGTCACATCAGGTATGCTTCTGCTCATTT	1727	CGCATCGGAGTGAAGAGCGCTTTGTTAGCAGCAGGACCTGGTATTCGGTGTGTGGAAACA	1786
DB	408 AGATGGCAATGAGACTGCTCCCTGAGGTGTGTCACATCAGGTATGCTTCTGCTCATTT	2009	CAGGGTCTCTCAGTGTATCTCTGGGGCGCTGGGCACTGATGACACAGAAAGTTAA	2068
QY	690 TTTTATGATGCTGCTTATAATTTGACTCGATTTAATATTAATTACAGTTTGTATGTG	1787	CAGGGTCTCTCAGTGTATCTCTGGGGCGCTGGGCACTGATGACACAGAAAGTTAA	1846
DB	468 TTTTATGATGCTGCTTATAATTTGACTCGATTTAATATTAATTACAGTTTGTATGTG	2069	AATCAGATGTTTTCCTGAGTGTGTTTCCAAAGAACTCTTGACCATGACAGATGTGACACAGAT	2128
QY	750 TCCAAATACCTGCTCAGCGCGAGGAGTGAAGTCAATAGTATAGCAGGAACTGTGA			
DB	528 TCCAAATACCTGCTCAGCGCGAGGAGTGAAGTCAATAGTATAGCAGGAACTGTGA			
QY	810 ATGTGATGTTCTGAAAACCTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA			
DB	588 ATGTGATGTTCTGAAAACCTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA			
QY	870 CTGTGTTTTCCTCATCGAGGCAATCTGCAATTCAGTGTATGTCAGAGGATGCTCTGCTT			
DB	648 CTGTGTTTTCCTCATCGAGGCAATCTGCAATTCAGTGTATGTCAGAGGATGCTCTGCTT			
QY	930 CTCAGACTGCGGGCTCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG			
DB	708 CTCAGACTGCGGGCTCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG			

Db 1847 AATCAGATGTTTTCCTCAAGAACTCTTGACCATGACAGATGTGACCAAGCACAGATT 1906  
Qy 2129 GTTACAGCTGTA CAGCAACACCAATGACTGCCACTGGTGCAATGACCAATTTGTGCCCA 2188  
Db 1907 GTTACAGCTGCACAGCAACACCAATGACTGCCACTGGTGCAATGACCAATTTGTGCCCA 1966  
Qy 2189 GGAACACAGCTGCTCAGAGGCGCATCTCCATTTTATAGTATGAGATTTGCCCAAGG 2248  
Db 1967 GGAACACAGCTGCTCAGAGGCGCATCTCCATTTTATAGTATGAGATTTGCCCAAGG 2026  
Qy 2249 ATAACCCCTATGCTACTGTAAACAAGACAGCTGCAAGGAGCTGTGCCCTGGACCCAGA 2308  
Db 2027 ATAACCCCTATGCTACTGTAAACAAGACAGCTGCAAGGAGCTGTGCCCTGGACCCAGA 2086  
Qy 2309 ACTGCCAGTGGGAGCCCGGAATCAGAGTGCAATTCGCCCTGCCGGAATATCTGTGGCA 2369  
Db 2087 ACTGCCAGTGGGAGCCCGGAATCAGAGTGCAATTCGCCCTGCCGGAATATCTGTGGCA 2146  
Qy 2369 TTGGCTGGCAATTTGGTTGGAATCTCATGTTTGAATAATTAATCTGCAAGGAGATTTATG 2428  
Db 2147 TTGGCTGGCAATTTGGTTGGAATCTCATGTTTGAATAATTAATCTGCAAGGAGATTTATG 2206  
Qy 2429 ACATGCTAAATTTGCTGTAGGAACCAATGCCCCCTTTGGCTTCTCTTACAAACCCAGA 2488  
Db 2207 ACATGCTAAATTTGCTGTAGGAACCAATGCCCCCTTTGGCTTCTCTTACAAACCCAGA 2266  
Qy 2489 AGAAGGTAGAAATTTGCTTTAGAGAGCTGCGAATATGCAATGCTGCAAGGAGATTTATG 2548  
Db 2267 AGAAGGTAGAAATTTGCTTTAGAGAGCTGCGAATATGCAATGCTGCAAGGAGATTTATG 2326  
Qy 2549 AGCTACCTTAAACCCCATGGTTCGCTTCGGAAGATCAATGCTGCTGCTGCTGGG 2608  
Db 2327 AGCTACCTTAAACCCCATGGTTCGCTTCGGAAGATCAATGCTGCTGCTGCTGGG 2386  
Qy 2609 AGATATGTCCTCATTAATGATTTACTACAGTGAATGCGCTGAGGCGGAGTGAATG 2668  
Db 2387 AGATATGTCCTCATTAATGATTTACTACAGTGAATGCGCTGAGGCGGAGTGAATG 2446  
Qy 2669 CTGGATTTCTGCGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2728  
Db 2447 CTGGATTTCTGCGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCA 2506  
Qy 2729 TCAACCCACTCAATGGTATGCTGTGAAGGCTGCAAAACAGTGTAAAGCAGTGC 2788  
Db 2507 TCAACCCACTCAATGGTATGCTGTGAAGGCTGCAAAACAGTGTAAAGCAGTGC 2566  
Qy 2789 GGACACCATGTCCTTGAAGGAGCAGCATGTGAGATTTGCACAGCGGAGCTGAGTGCA 2848  
Db 2567 GGACACCATGTCCTTGAAGGAGCAGCATGTGAGATTTGCACAGCGGAGCTGAGTGCA 2626  
Qy 2849 TGTGTGAGCAACATGAAGCAGTGTGGAATCTCAATGCTGCTGCTGCTTCCCTT 2908  
Db 2627 TGTGTGAGCAACATGAAGCAGTGTGGAATCTCAATGCTGCTGCTGCTTCCCTT 2686  
Qy 2909 TTGGCCAGTGTGAATGATATGATGAGCAGTGCCTGCCCCCTGAAATTTGTCAGCT 2968  
Db 2687 TTGGCCAGTGTGAATGATATGATGAGCAGTGCCTGCCCCCTGAAATTTGTCAGCT 2746  
Qy 2969 ACTGTACCTGTAGTCAATGCTTGGAGCAACAGAGCTGTGGCTGTACTGATCCCAAGCA 3028  
Db 2747 ACTGTACCTGTAGTCAATGCTTGGAGCAACAGAGCTGTGGCTGTACTGATCCCAAGCA 2806  
Qy 3029 ATACTGGCAAGGGAATGCAATAGAGGTTCTTATTAAGGACAGTGAAGATGCTTCGC 3088  
Db 2807 ATACTGGCAAGGGAATGCAATAGAGGTTCTTATTAAGGACAGTGAAGATGCTTCGC 2866  
Qy 3089 AAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATTCAGCATGTGCTAGAGG 3148  
Db 2867 AAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATTCAGCATGTGCTAGAGG 2926  
Qy 3149 ACAGAGATACATGCTGCTTCTTATTCATGCTGAGCTTGCATGCAACCGGCCACAGTA 3208

Db 2927 ACAGAGATACATGCTGCTTCTTCAATTCAGTGTCCAGTTCGCAATGCAACGGCCACAGTA 2986  
Qy 3209 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTTGACACAGGCAAGCAGCTGCG 3268  
Db 2987 AATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTTGACACAGGCAAGCAGCTGCG 3046  
Qy 3269 AGACTGCAATCTGGCTTCTAGCGTATCCCAACCAATGAGGGAAATGTCAGCCATGCA 3328  
Db 3047 AGACTGCAATCTGGCTTCTAGCGTATCCCAACCAATGAGGGAAATGTCAGCCATGCA 3106  
Qy 3329 AGTCAATGGGACGCTCTCTGTGCAACCAACAGGCAAGTCTCTGCAACCA 3388  
Db 3107 AGTCAATGGGACGCTCTCTGTGCAACCAACAGGCAAGTCTCTGCAACCA 3166  
Qy 3389 AGGCGTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3448  
Db 3167 AGGCGTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACC 3226  
Qy 3449 CTCTCAGAGCAATGTTATTTATTAATCTCTTCTTATGACTATCAGTTTCACTTTAGTCTAT 3508  
Db 3227 CTCTCAGAGCAATGTTATTTATTAATCTCTTCTTATGACTATCAGTTTCACTTTAGTCTAT 3286  
Qy 3509 CCCAGGAATGATGCTGCTATTTACACAGCTATCAATTTTGTGGCTACTCTCTGACGAACAAA 3568  
Db 3287 CCCAGGAATGATGCTGCTATTTACACAGCTATCAATTTTGTGGCTACTCTCTGACGAACAAA 3346  
Qy 3569 ACAGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTTCAACCTCAACATCACCTGGG 3628  
Db 3347 ACAGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTTCAACCTCAACATCACCTGGG 3406  
Qy 3629 CTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAACCA 3688  
Db 3407 CTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAACCA 3466  
Qy 3689 ACATTAAGAGTACAAAGATAGTTCCTTAATGAGAAGTTGATTTGCAACACCCCAA 3748  
Db 3467 ACATTAAGAGTACAAAGATAGTTCCTTAATGAGAAGTTGATTTGCAACACCCCAA 3526  
Qy 3749 ATATCACTTTCTTTGTTTATGTCAGTAATTTCACTGGCCCATCAAAATTCAGTGCATA 3808  
Db 3527 ATATCACTTTCTTTGTTTATGTCAGTAATTTCACTGGCCCATCAAAATTCAGTGCATA 3586  
Qy 3809 CTGAACAATGA 3819  
Db 3587 CTGAACAATGA 3597

## RESULT 8

AAS72657  
ID AAS72657 standard; cDNA; 3597 BP.

XX AAS72657;

XX AC

XX XX

DT 13-FEB-2002 (first entry)

XX

DE DNA encoding novel human diagnostic protein #8461.

XX

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

OS Homo sapiens.

XX

PN W0200175067-A2.

XX

PD 11-OCT-2001.

XX

XX 30-MAR-2001; 2001WO-US008631.

XX

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00649167.

XX

PA (HYSB-) HYSEQ INC.

XX





Db 1607 TGAATGCAACAGATTTGGGCAATTCAGAGCTCTTACACAGCAGCCATGATGATGTTGG 1666  
Qy 1889 GTGGTTTCAATAGTCTCTCTCAGGACATCTCGTATTTACCTCGGAACAGTGTGATG 1948  
Db 1667 GTGGTTTCAATAGTCTCTCTCAGGACATCTCGTATTTACCTCGGAACAGTGTGATG 1726  
Qy 1949 CGCATCGGAGTGAAGCCGCTTGTATAGCAGCAGGACCTCGTATTTGGTGTGTGTGAACA 2008  
Db 1727 CGCATCGGAGTGAAGCCGCTTGTATAGCAGCAGGACCTCGTATTTGGTGTGTGTGAACA 1786  
Qy 2009 CAGGTCCTCTCAGTGTATCTCGTGGCGCTGCAACTGATGAACAAGAGAAAAGTTAA 2068  
Db 1787 CAGGTCCTCTCAGTGTATCTCGTGGCGCTGCAACTGATGAACAAGAGAAAAGTTAA 1846  
Qy 2069 AATCAGATGTTTTTCCAAAAGAACTCTTGAACATGACAGATGTGACCAAGCACAGATT 2128  
Db 1847 AATCAGATGTTTTTCCAAAAGAACTCTTGAACATGACAGATGTGACCAAGCACAGATT 1906  
Qy 2129 GTTACAGCTGTACAGCCACCAATGACTGCACTGGTCAATGACCAATGTTGTCCTCA 2188  
Db 1907 GTTACAGCTGTACAGCCACCAATGACTGCACTGGTCAATGACCAATGTTGTCCTCA 1966  
Qy 2189 GGAACACACAGCTGCTCAGAGGCCAGATCTCCATTTTGTAGTATGAGAAATTCGCCCAAGG 2248  
Db 1967 GGAACACACAGCTGCTCAGAGGCCAGATCTCCATTTTGTAGTATGAGAAATTCGCCCAAGG 2026  
Qy 2249 ATAAACCCATGTACTACTGTAAACAGAGACAGCTGCAAGAGCTGTGCCCTGGACACAGA 2308  
Db 2027 ATAAACCCATGTACTACTGTAAACAGAGACAGCTGCAAGAGCTGTGCCCTGGACACAGA 2086  
Qy 2309 ACTGCCAGTGGAGCCCGGAACTCAGAGTGCATTTGCCCTGCCGAAATATCTGTGCA 2368  
Db 2087 ACTGCCAGTGGAGCCCGGAACTCAGAGTGCATTTGCCCTGCCGAAATATCTGTGCA 2146  
Qy 2369 TTGGCTGGCAATTTGGTTGAAACTCATGTTGAAATTTACTACTGCCAAGGAGAAATATG 2428  
Db 2147 TTGGCTGGCAATTTGGTTGAAACTCATGTTGAAATTTACTACTGCCAAGGAGAAATATG 2206  
Qy 2429 ACAATGCTAAATTTGTTCTGAGAACCAATGCTGCTTGGCTTCTCTTCAACCCAGA 2488  
Db 2207 ACAATGCTAAATTTGTTCTGAGAACCAATGCTGCTTGGCTTCTCTTCAACCCAGA 2266  
Qy 2489 AGAAGTAGAATTTGTTCTTAAAGCAGCTGCGAATATGCAATGCTCAGAGCATGTCCA 2548  
Db 2267 AGAAGTAGAATTTGTTCTTAAAGCAGCTGCGAATATGCAATGCTCAGAGCATGTCCA 2326  
Qy 2549 AGCTACCTTAACCCATGGTGGCTTCGGAAGATCAATGTGCTTACTGGTGGTGG 2608  
Db 2327 AGCTACCTTAACCCATGGTGGCTTCGGAAGATCAATGTGCTTACTGGTGGTGG 2386  
Qy 2609 AAGATATGTCCTCAATTTACAAATAGTTTACTACAGTGCCTGCTCAGGCCAGTGATG 2668  
Db 2387 AAGATATGTCCTCAATTTACAAATAGTTTACTACAGTGCCTGCTCAGGCCAGTGATG 2446  
Qy 2669 CTGGATTCGTGGAATTTTATCAGAACCCAGTACTCGGGGAGTGAAGCTGCAACCTGCA 2728  
Db 2447 CTGGATTCGTGGAATTTTATCAGAACCCAGTACTCGGGGAGTGAAGCTGCAACCTGCA 2506  
Qy 2729 TCAACCCACTCAATGCTAGTGTCTGTGAAGCCCTGCAAAACACAGTCTAAGCAGTGCC 2788  
Db 2507 TCAACCCACTCAATGCTAGTGTCTGTGAAGCCCTGCAAAACACAGTCTAAGCAGTGCC 2566  
Qy 2789 GGACACCATGTCCCTTGAAGACAGCATGTGGAGATTGACACAGCGGACGCTCTGAGTGA 2848  
Db 2567 GGACACCATGTCCCTTGAAGACAGCATGTGGAGATTGACACAGCGGACGCTCTGAGTGA 2626  
Qy 2849 TGTGGTGCAGCAACATGAAGCAGTGTGGACTCCCAATGCTTATGTGGCTCTCCCTT 2908  
Db 2627 TGTGGTGCAGCAACATGAAGCAGTGTGGACTCCCAATGCTTATGTGGCTCTCCCTT 2686  
Qy 2909 TTGGCCAGTGTATGGAATGTTATGATGAGCAGCAGCTGCCCTGCCCTGAAATTTGTCAGGCT 2968  
Db 2687 TTGGCCAGTGTATGGAATGTTATGATGAGCAGCAGCTGCCCTGCCCTGAAATTTGTCAGGCT 2746

Qy 2969 ACTGTACTGTAGTCAATTTGCTTGGAGCAACAGGCTGTGGCTGTGTACTGATCCAGCA 3028  
Db 2747 ACTGTACTGTAGTCAATTTGCTTGGAGCAACAGGCTGTGGCTGTGTACTGATCCAGCA 2806  
Qy 3029 ATACTGCAAGAGGAATGATAGAGGTTCTTATAAGAGACAGTGAAGATCCCTTCG 3088  
Db 2807 ATACTGCAAGAGGAATGATAGAGGTTCTTATAAGAGACAGTGAAGATCCCTTCG 2866  
Qy 3089 AAGCCCTACAGAAATTTCTATCCACAGCCCTGTCTCAATTCAGCATGTGTAGAGG 3148  
Db 2867 AAGCCCTACAGAAATTTCTATCCACAGCCCTGTCTCAATTCAGCATGTGTAGAGG 2926  
Qy 3149 ACAGCAGATCAACTGTGCTTTTCAATTCATCTGTCAGCTTCCCAATGCAAGCCACAGTA 3208  
Db 2927 ACAGCAGATCAACTGTGCTTTTCAATTCATCTGTCAGCTTCCCAATGCAAGCCACAGTA 2986  
Qy 3209 AATGCAATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACACAGGCAAGCACTGG 3268  
Db 2987 AATGCAATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACACAGGCAAGCACTGG 3046  
Qy 3269 AGACCTGCATATCTGGCTTCTACGGTATCCCAACCAATGGAGGAAATGTCAGCCATGA 3328  
Db 3047 AGACCTGCATATCTGGCTTCTACGGTATCCCAACCAATGGAGGAAATGTCAGCCATGA 3106  
Qy 3329 AGTGCANTGGCAGCGCTCTCTGTGCAACACCAACAGCGGCAAGTCTTCTGCAACCA 3388  
Db 3107 AGTGCANTGGCAGCGCTCTCTGTGCAACACCAACAGCGGCAAGTCTTCTGCAACCA 3166  
Qy 3389 AGGCGGTCAAGGGGACAGTGCAGCTATGTGAGGTAGAAAATCGATATCAAGGAAAC 3448  
Db 3167 AGGCGGTCAAGGGGACAGTGCAGCTATGTGAGGTAGAAAATCGATATCAAGGAAAC 3226  
Qy 3449 CTCTCAGAGAAATGTTATTTACTCTTCTTATTGACTATCAGTTTACCTTTAGTCTAT 3508  
Db 3227 CTCTCAGAGAAATGTTATTTACTCTTCTTATTGACTATCAGTTTACCTTTAGTCTAT 3286  
Qy 3509 CCAGGAAGATGATCGTATTACACAGCTATCAATTTTGTGGCTACTCTCTGACGAACAA 3568  
Db 3287 CCAGGAAGATGATCGTATTACACAGCTATCAATTTTGTGGCTACTCTCTGACGAACAA 3346  
Qy 3569 ACAGGATTTGGACATGTTTCAATTCCTTCCAAAGATTTCAACCTCAACATCACCTGG 3628  
Db 3347 ACAGGATTTGGACATGTTTCAATTCCTTCCAAAGATTTCAACCTCAACATCACCTGG 3406  
Qy 3629 CTGCCATTTCTCAGCTGGNAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCA 3688  
Db 3407 CTGCCATTTCTCAGCTGGNAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCA 3466  
Qy 3689 ACATTAAGGAGTACAAAGATGTTTCTTAATGAGAAGTTTGTATTTTCGCAACCCCAA 3748  
Db 3467 ACATTAAGGAGTACAAAGATGTTTCTTAATGAGAAGTTTGTATTTTCGCAACCCCAA 3526  
Qy 3749 ATATCATTTTCTTGTATTATGTCAGTAATTTCACTGGGCCATCAAAATTCAGTGCAAA 3808  
Db 3527 ATATCATTTTCTTGTATTATGTCAGTAATTTCACTGGGCCATCAAAATTCAGTGCAAA 3586  
Qy 3809 CTGAACAATGA 3819  
Db 3587 CTGAACAATGA 3597

RESULT 9

ADD70997

ID ADD70997 standard; DNA; 3597 BP.

XX ADD70997;

AC

XX

Dr 15-JAN-2004 (first entry)

XX

DE Human attractin gene SEQ ID NO:1.

XX

KW liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;



KW cytostatic; gene therapy; human; gene; ds.  
XX Homo sapiens.  
OS WO2003061564-A2.  
XX 31-JUL-2003.  
XX 20-DEC-2002; 2002WO-US040718.  
XX 21-DEC-2001; 2001US-0341815P.  
XX 31-DEC-2001; 2001US-0343185P.  
XX (GENE-) GENE LOGIC INC.  
PA (LGBI-) LG BIOMEDICAL INST.  
XX Koh SS, Liu Q, Chung H, Zeng W, Lee B, Yeramilli S, Song SY;  
PI WPI; 2003-663343/62.  
XX Diagnosing liver cancer cells, useful for treating liver cancer.  
PT associated with chronic hepatitis or cirrhosis comprises detecting the  
PT level of expression in a tissue sample of one or more genes associated  
PT with cancerous liver tissues.  
XX Claim 1; SEQ ID NO 1; 176pp; English.  
XX The present invention describes a method for diagnosing liver cancer  
CC cells comprising detecting the level of expression in a tissue sample of  
CC one or more genes given in the specification (see ABD70997 to ABD71105),  
CC where differential expression of the genes is indicative of liver cancer.  
CC Also described: (1) detecting the progression of liver cancer in a  
CC patient; (2) monitoring the treatment of a patient with liver cancer; (3)  
CC treating a patient with liver cancer; (4) typing a liver disease in a  
CC patient; (5) detecting the presence or progression of liver cancer in a  
CC patient with chronic hepatitis or cirrhosis; (6) differentiating liver  
CC cancer related to chronic hepatitis from liver cancer related to  
CC cirrhosis; (7) screening for an agent capable of modulating the onset or  
CC progression of liver cancer; (8) a composition comprising at least two  
CC oligonucleotides comprising a sequence that specifically hybridises to  
CC any of the genes; (9) a solid support comprising the at least two  
CC oligonucleotides; (10) a computer system comprising a database containing  
CC information identifying the level in liver tissue of a set of genes; (11)  
CC a method for using the computer system to present information identifying  
CC the expression level in tissue or cell of any of the genes; and (12) a  
CC therapeutic agent for slowing or halting the progression of liver cancer.  
CC The methods are useful for treating liver cancer associated with chronic  
CC hepatitis or cirrhosis. The present sequence represents a specifically  
CC claimed human gene sequence which is used in the exemplification of the  
CC present invention.  
XX SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;  
Query Match 91.3%; Score 3487; DB 9; Length 3597;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 3524; Conservative 0; Mismatches 25; Indels 2; Gaps 2;  
QY 270 GTCGGCTCAGCGGACGCGAGCGAAGGAATGACCGGCCCTCTGTCAACGGCGGTGCG 329  
DB 48 GCGGGCGACGGCAGCGCTCGCGGGCAGGAGCGCGGGCGGCACTGTGTCAACGGCGGTGCG 107  
QY 330 CTGCAACCTGGCCACCGCCAGTGGCTGTGCCCCCGCGGCTGGGTGGCGGAGCAATGCCA 389  
DB 108 CTGCAACCTGGCCACCGCCAGTGGCTGTGCCCCCGCGGCTGGGTGGCGGAGCAATGCCA 167  
QY 390 GCACTCGCGGGCGGCTTCAGACTAATCTGATCTTCTGGTGTGTGACAGATGGACCTGG 449  
DB 168 GCACTCGCGGGCGGCTTCAGACTAATCTGATCTTCTGGTGTGTGACAGATGGACCTGG 227  
QY 450 AAATATAAATAACAAACGAAGTGACAGTGGCTCAATGGAAGCAGCCAAATAGAAATAAT 509  
DB 228 AAATATAAATAACAAACGAAGTGACAGTGGCTCAATGGAAGCAGCCAAATAGAAATAAT 287

510 GAGACTTCGTTTCAATCATTTTCTACAGAGTGTAGTTGGGACCACTTTATATATGTTATGA 569  
288 GAGACTTCGTTTCAATCATTTTCTACAGAGTGTAGTTGGGACCACTTTATATATGTTATGA 347  
570 TGGGACTCAATTTATGACACCGCTAGTTCTGCTGCAITTTAGTGGCCTCATTTTCTCTGAGAG 629  
348 TGGGACTCAATTTATGACACCGCTAGTTCTGCTGCAITTTAGTGGCCTCATTTTCTCTGAGAG 407  
630 AGATGCAATGAGACTGTCCCTGAGTGTGTCGACATCAGATGATGCTTCTGCTGCAITTT 589  
408 AGATGCAATGAGACTGTCCCTGAGTGTGTCGACATCAGATGATGCTTCTGCTGCAITTT 467  
690 TTTTAGTGTGCTGCTTATAATTTGACTCGATTATATATTTACTTACAGTTTTGATATGTG 749  
468 TTTTAGTGTGCTGCTTATAATTTGACTCGATTATATATTTACTTACAGTTTTGATATGTG 527  
750 TCCAAATAACTGCTCAGCGGAGAGTGTAGATCAGTAAATAGCAGGAACTGTTGA 809  
528 TCCAAATAACTGCTCAGCGGAGAGTGTAGATCAGTAAATAGCAGGATGCTTCTGCTTGA 587  
810 ATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACACACAA 869  
588 ATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACACACAA 647  
870 CTGTGTTTCTCTCATCGAGGATCTGCAATTCAGTGTGTCAGAGGATGCTTCTGCTT 929  
648 CTGTGTTTCTCTCATCGAGGATCTGCAATTCAGTGTGTCAGAGGATGCTTCTGCTT 707  
930 CTGAGACTGCGAGGTCCTGGATGTTTCAGTTCCTGTACCAGCTAAACAGTCAATTTGGAC 989  
708 CTGAGACTGCGAGGTCCTGGATGTTTCAGTTCCTGTACCAGCTAAACAGTCAATTTGGAC 767  
990 TCCAGAGGAATATTTCTAACTTAAAGTCTCCAGAGCATCTCAATAAGCTGTGCTCAATGG 1049  
768 TCCAGAGGAATATTTCTAACTTAAAGTCTCCAGAGCATCTCAATAAGCTGTGCTCAATGG 827  
1050 AAACATATGTTGGTGTGTCGAGATATATGTTCAACACTCAGATTAACATGTTCT 1109  
828 AAACATATGTTGGTGTGTCGAGATATATGTTCAACACTCAGATTAACATGTTCT 887  
1110 AGCGTATGACTTCTTACGAGGAGTGGCTTCCACTAAACCGTCTGTGAAACATGTTGT 1169  
888 AGCGTATGACTTCTTACGAGGAGTGGCTTCCACTAAACCGTCTGTGAAACATGTTGT 947  
1170 TGTAGATATGTCATTTCTTTGGCATTTATAACAGGATAAAATTTACATGATGGAGAA 1229  
948 TGTAGATATGTCATTTCTTTGGCATTTATAACAGGATAAAATTTACATGATGGAGAA 1007  
1230 AATTGATCAACTGGGATGTGACCAATGATGAGTTGAGAGTTTTTTCATATTCATATGATC 1289  
1008 AATTGATCAACTGGGATGTGACCAATGATGAGTTGAGAGTTTTTTCATATTCATATGATC 1067  
1290 ATGGGTGTTGTTGACCCCTTAAGCAAGGAGCAGTATGCAAGTGGTTGGGCACTCTGCACA 1349  
1068 ATGGGTGTTGTTGACCCCTTAAGCAAGGAGCAGTATGCAAGTGGTTGGGCACTCTGCACA 1127  
1350 CATTTGTACACTGAGAAATGCGGAGTGGTGTGCTGCTCATCTTTGGTCACTGCCCTCT 1409  
1128 CATTTGTACACTGAGAAATGCGGAGTGGTGTGCTGCTCATCTTTGGTCACTGCCCTCT 1187  
1410 CTATGGATATATAGCAATGTGCAAGGAATATGATTTTGGATAAGAAACACATGGAGTATATT 1469  
1188 CTATGGATATATAGCAATGTGCAAGGAATATGATTTTGGATAAGAAACACATGGAGTATATT 1247  
1470 ACACACCCAGGGTGCCTTTGTGAAGGGGTTACGGCCATAGCAGTGTTCACACCAATAG 1529  
1248 ACACACCCAGGGTGCCTTTGTGAAGGGGTTACGGCCATAGCAGTGTTCACACCAATAG 1307  
1530 GACACGGCCCTTACCTTCACTGTTGCTACAGGCTTTTTCAGTGCCCAATAGTACCGGCT 1589  
1308 GACACGGCCCTTACCTTCACTGTTGCTACAGGCTTTTTCAGTGCCCAATAGTACCGGCT 1367  
1590 TGCAGATGATCTCTACCGGATATGATGTGGATATCCAGATGTGGACCACTTCTTAAGGACAG 1649

Db	1368	TTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCAATCTTTAAGGACAG	1427
Qy	1650	CCGATTTTTTCGGTTACTTTGGCAACAGAGCTGTGATAGTGGAGAACCAATGCTGGTGTTTGG	1709
Db	1428	CCGATTTTTTCGGTTACTTTGGCAACAGAGCTGTGATAGTGGAGAACCAATGCTGGTGTTTGG	1487
Qy	1710	GGGAACAACACACAATATGACATCTATGAGCGCATGGGCCCAAAATGCTTCTTTCAGATTT	1769
Db	1488	GGGAACAACACACAATATGACATCTATGAGCGCATGGGCCCAAAATGCTTCTTTCAGATTT	1547
Qy	1770	CATGGCCTATGACATTTGCTGTGACCGCTGTGTCAAGTCTTCCACAGACTGATCTCCACCA	1829
Db	1548	CATGGCCTATGACATTTGCTGTGACCGCTGTGTCAAGTCTTCCACAGACTGAT - TCCACCA	1606
Qy	1830	TGATGTCAACAGATTTGGCCATT - CAGCAGCTTTTACACAACAGCACCATATGATGTTCG	1888
Db	1607	TGATGTCAACAGATTTGGCCATTCCAGCAGCTTTTACACAACAGCACCATATGATGTTCG	1666
Qy	1889	GTGGTTCAATAGTACTCTCTCAGCGACATCTGTGATTTACACTCGGACAGTGTGATG	1948
Db	1667	GTGGTTTCAAATAGTCTCTCTCAGCGACATCTGTGATTTACACTCGGACAGTGTGATG	1726
Qy	1949	CGCATCGGAGTGAAGCCGCTTTGTTTAGCAGCAGCAGCTGTGATTTGCGTGTGTGGAAACA	2008
Db	1727	CGCATCGGAGTGAAGCCGCTTTGTTTAGCAGCAGCAGCCTGTGATTTGCGTGTGTGGAAACA	1786
Qy	2009	CAGGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAAACAGAGAAGAAAGTTAA	2068
Db	1787	CAGGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAAACAGAGAAGAAAGTTAA	1846
Qy	2069	AATCAGAAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACACACACAGATT	2128
Db	1847	AATCAGAAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACACACACAGATT	1906
Qy	2129	GTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTGCATATGACCAATGTGTCCCCA	2188
Db	1907	GTTACAGCTGTACAGCCAAACCAATGACTGCCACTGGTGCATATGACCAATGTGTCCCCA	1966
Qy	2189	GGAAACACAGCTGCTCAGAGGCCAGATCTCCATTTTTAGTGTATGAGAAATGGCCCAAGG	2248
Db	1967	GGAAACACAGCTGCTCAGAGGCCAGATCTCCATTTTTAGTGTATGAGAAATGGCCCAAGG	2026
Qy	2249	ATAACCCCATGTACTACTGTAAACAGAGAACCCAGCTGCAGGAGCTGTGCCCTGGACCCAGA	2308
Db	2027	ATAACCCCATGTACTACTGTAAACAGAGAACCCAGCTGCAGGAGCTGTGCCCTGGACCCAGA	2086
Qy	2309	ACTGCCAGTGGAGCCCGGATCAGCAGGTGCATTTGCCCTGCCGGAATAATCTGTGGCA	2368
Db	2087	ACTGCCAGTGGAGCCCGGATCAGCAGGTGCATTTGCCCTGCCGGAATAATCTGTGGCA	2146
Qy	2369	TTGGCTGGCATTTGGTTGGAAATCATGTGTTGAAAATTAATACTGCCCAAGGAGAAATATG	2428
Db	2147	TTGGCTGGCATTTGGTTGGAAATCATGTGTTGAAAATTAATACTGCCCAAGGAGAAATATG	2206
Qy	2429	ACAAATGCTAAATTTGTTCTGTATAGGAACCAATATGCCCTTTTGGCTTCTTTACAACCCAGA	2488
Db	2207	ACAAATGCTAAATTTGTTCTGTATAGGAACCAATATGCCCTTTTGGCTTCTTTACAACCCAGA	2266
Qy	2489	AGAAGGTAGAAATTTGTCCTTAAGCAGCTGCGAATAATGCAAGTCAATCTCAGAGCATGTCCA	2548
Db	2267	AGAAGGTAGAAATTTGTCCTTAAGCAGCTGCGAATAATGCAAGTCAATCTCAGAGCATGTCCA	2326
Qy	2549	AGTCACTTAAACCCCAATGGGTGGCCCTTCGGAAGATCAAATGTGTCTACTGTGTGGG	2608
Db	2327	AGTCACTTAAACCCCAATGGGTGGCCCTTCGGAAGATCAAATGTGTCTACTGTGTGGG	2386
Qy	2609	AAGATATGCCCCATTTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAAGTATG	2668
Db	2387	AAGATATGCCCCATTTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAAGTATG	2446
Qy	2669	CTGGATTTCTGTGGAATTTTATCAGAACCCCAAGTACTCGGGAGCTCAAGGCTGCAACCTGCA	2728

QY 3809 CTGACAAATGA 3819  
DB 3587 CTGACAAATGA 3597

RESULT 10  
ADBS3249  
ID ADBS3249 standard; DNA; 8739 BP.  
AC ADBS3249;  
XX  
XX  
DT 04-DEC-2003 (first entry)  
XX  
XX  
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:3791.  
XX  
XX  
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;  
XX toxicity marker; toxicity progression; drug screening;  
KW primary rat hepatocyte toxicity modelling; gene; ds.  
XX  
XX  
OS Rattus norvegicus.  
XX  
XX  
PN WO2003065993-A2.  
XX  
XX  
PD 14-AUG-2003.  
XX  
XX  
PF 04-FEB-2003; 2003WO-US003482.  
XX  
XX  
PR 04-FEB-2002; 2002US-0353171P.  
PR 13-MAR-2002; 2002US-0363534P.  
PR 08-APR-2002; 2002US-0370248P.  
PR 10-APR-2002; 2002US-0371134P.  
PR 10-APR-2002; 2002US-0371135P.  
PR 10-APR-2002; 2002US-0371150P.  
PR 11-APR-2002; 2002US-0371413P.  
PR 19-APR-2002; 2002US-0373601P.  
PR 19-APR-2002; 2002US-0373602P.  
PR 22-APR-2002; 2002US-0374139P.  
PR 08-MAY-2002; 2002US-0378370P.  
PR 09-MAY-2002; 2002US-0378652P.  
PR 09-MAY-2002; 2002US-0378653P.  
PR 09-MAY-2002; 2002US-0378655P.  
PR 09-JUL-2002; 2002US-0394230P.  
PR 09-JUL-2002; 2002US-0394253P.  
PR 04-SEP-2002; 2002US-0407688P.  
PR 28-JAN-2003; 2003US-0442900P.  
XX  
XX  
PA (GENE-) GENE LOGIC INC.  
XX  
XX  
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;  
PI Elashoff M;  
XX  
XX  
DR WPI; 2003-731472/69.  
XX  
XX  
PT Determining if a compound induces a toxic effect on a tissue or cell, for  
PT identifying hepatotoxic compounds, comprises comparing a gene expression  
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox  
PT mean values.  
XX  
XX  
PS Claim 44; SEQ ID NO 3791; 874pp; English.  
XX  
XX  
CC The present invention describes a method for determining whether a  
CC compound induces a toxic effect on a tissue or cell. The method comprises  
CC preparing a gene expression profile of a tissue or cell sample exposed to  
CC the compound, and comparing the gene expression profile to a database  
CC comprising data or information on the Tox mean and non-Tox mean value.  
CC The method is useful for predicting or identifying at least one toxic  
CC effect, particularly hepatotoxicity, of a test or unknown compound. The  
CC genes listed in the specification are useful as diagnostic or toxicity  
CC markers for the prediction or identification of the physiological state  
CC of tissue or cell sample that has been exposed to a compound, or to  
CC identify or predict the toxic effects of a compound or an agent. These  
CC may also be used as markers for monitoring toxicity progression or for

CC drug screening. The present sequence represents a primary rat hepatocyte  
CC toxicity modelling related gene sequence from the present invention.  
XX  
XX  
SQ Sequence 8739 BP; 2288 A; 2115 C; 2109 G; 2227 T; 0 U; 0 Other;  
Query Match 81.8%; Score 3114.8; DB 9; Length 8739;  
Best Local Similarity 88.8%; Pred. No. 0;  
Matches 3371; Conservative 0; Mismatches 427; Indels 0; Gaps 0;  
QY 3 GGTGGCGGAGCGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACCGCGCGACCGCAGC 62  
DB 72 GCGCGGCGGCGGAGCGGACTGAGCGCGGCTGAGGGGGTACACGACGCGACACGAC 131  
QY 63 GCTCGCGGCGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 122  
DB 132 GCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 191  
QY 123 GGGGCTGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 182  
DB 192 GTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 251  
QY 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242  
DB 252 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 311  
QY 243 CGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 302  
DB 312 CGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 371  
QY 303 TGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 362  
DB 372 TGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 431  
QY 363 CGCGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 422  
DB 432 CACGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 491  
QY 423 TTCTGGGTTTGTACAGATGCGACCTGGAATTAATAATACAAACCAAGTGACGCTGG 482  
DB 492 TTCTGGGTTTGTACAGATGCGACCTGGAATTAATAATACAAACCAAGTGACGCTGG 551  
QY 483 CATTGAAGGACGCGCAAAATAGAAATATAGACTTCGTTTCAATTCATTTTGTCTACAG 542  
DB 552 CATTGAAGGACGCGCAAAATAGAAATATAGACTTCGTTTCAATTCATTTTGTCTACAG 611  
QY 543 TAGTTGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC 602  
DB 612 TAGTTGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGC 671  
QY 603 ATTTAGTGGCCTCATTTGTTCTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTG 662  
DB 672 CTTTAGTGGCCTCATTTGTTCTTCTGAGAGAGATGGCAATGAGAGCGGCCCTGAGTCACTGT 731  
QY 663 CACATCAGGTTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 722  
DB 732 CACTTCAGGTTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791  
QY 723 TAATATTACTTACAGTTTTCATATGTTTCCAAATAACTGCTCAGGCGGAGAGAGTAA 782  
DB 792 TAATATTACTTACAGTTTTCATATGTTTCCCAATATGCTCAGGCGGAGAGAGTAA 851  
QY 783 GATCAGTAATAGCAGCGAAACTGTTGAATGTAATGTTTCTGAAACTGGAAGAGTGAAGC 842  
DB 852 GAGCAGTAACAGCAGCAGCAGCAGTGTGAGTGTGAATGTTTCTGAAACTGGAAGAGGAGTC 911  
QY 843 ATGTCGATTTCTCCTCAGTACAGACACTGCTGTTTCTCCTCAGGAGGCTGCAATTC 902  
DB 912 CTGTGACATTTCTCCTCAGTACAGACACTGCTGTTTCTCCTCAGGAGGCTGCAATTC 971  
QY 903 AAGTGATGTCAGAGGATGCTCCTGCTTCTCAGACTGCGAGGCTCCTGGATGTTCACTTC 962  
DB 972 AAGTGACACAGAGGCTGCTCCTGCTTCTCCTCAGTGGGCTCCTGGATGTTCAATTC 1031

QY 963 TGTACAGCTAACAGTCTATTTGGACTCGAGAGGATATTTCTAACTTAAAGCTCCCCAG 1022  
Db 1032 TGTGCCAGCTAACAGTCTATTTGGACTCGAGAGGATATTTCTAAAGCTCCCCAG 1091  
QY 1023 AGCATCTCATAAGCTGTGGTCAATGGAAATATATGTGGTGTGGAGGATATATGT 1082  
Db 1092 AGCTCTCATAAGCTGTGGTCAATGGAAATATATGTGGTGTGGAGGATATATGT 1151  
QY 1083 CAACCACTCAGATATTAACATGGTCTAGCGATGACCTTGTCTAGGAGTGGCTTCC 1142  
Db 1152 CAACCACTCAGATATTAACATGGTCTAGCGATGACCTTGTCTAGGAGTGGCTTCC 1211  
QY 1143 ACTAAACCGTCTGTGAAACAAATGTGGTGTAGATATGTCTATTTGGCAATATACAA 1202  
Db 1212 ACTAAACCAATTCTGTGAACAGTGTGGTGTAGATATGTCTATTTGGCAATATACAA 1271  
QY 1203 GGATAAAATTTACATGTATGAGGAGAAATATGATCCAACTGGGAATGTGACCAATGAGTT 1262  
Db 1272 GGATAAAATTTACATGTATGAGGAGAAATATGATCCAACTGGGAATGTGACCAAGCT 1331  
QY 1263 GAGAGTTTTTACATTCATAATGAGTCAATGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1322  
Db 1332 GAGAGTTTTTACATTCATAATGAGTCAATGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1391  
QY 1323 GTATGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1382  
Db 1392 GTATGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1451  
QY 1383 GCTGTGATCTTTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1442  
Db 1452 GTTGTGATCTTTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1511  
QY 1443 TTTGGTAAGAACACATGAGTATATTAACACACCGAGGTGGTGTGGTGTGGTGTGGTGTGGT 1502  
Db 1512 CTTGGTAAGAACACATGAGTATATTAACACACCGAGGTGGTGTGGTGTGGTGTGGTGTGGT 1571  
QY 1503 CGGCAATGAGCTGTATGACCAATGAGCAGGCGCTTACAGTGTATGAGTGTATGAGTGTATG 1562  
Db 1572 TGGCCACAGCAGTGTATGATGACAGGACCAAGGCTCTGTATGTTTCAATGGTGGCTACAA 1631  
QY 1563 GCTTTTCAAGTCAATGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1622  
Db 1632 GCTTTTCAAGTCAATGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1691  
QY 1623 CCAGATGTGGACCAATTTTAAAGACAGCGGATTTTCCGTTTACATGACACAGCTGTGAT 1682  
Db 1692 TCAGATGTGGACCAATTTTAAAGACAGCGGATTTTCCGTTTACATGACACAGCTGTGAT 1751  
QY 1683 AGTGAGTGGACCAATGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1742  
Db 1752 AGTGAGTGGACCAATGAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1811  
QY 1743 TGGCGCCAAATGCTTCTCTCAGATTTTCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1802  
Db 1812 CGGTGCCAAATGCTTCTCTCAGATTTTCAAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1871  
QY 1803 AGTGCTTCCAGACCTGATCTCCACCAATGATGTCAAGATTTGGCCATTTGACAGCTT 1862  
Db 1872 AGTTCTCCCAAGCTGAGCTCCATCAGATGTCAACCGATTTGGCCATTTGACAGCTT 1931  
QY 1863 ACACAAACAGCAGTATGTTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1922  
Db 1932 GCACAAACAGCAGTATGTTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGT 1991  
QY 1923 GGTATTCACCTCGGAAAGTGTGATGCGCATGCGAGTGAAGCCGCTGTTTGTAGCAGG 1982  
Db 1992 AGTCTTCACCTCGGAGAGTGTGATGACACCGCAGGAGCTGCTTGTGTGGCAGCAGG 2051  
QY 1983 ACCTGGTATGCGTGTGTGGAAACACAGGCTGTGCTCAGTGTATCTCGTGGCGCTGGC 2042  
Db 2052 ACCTGGCATCCGGTGTGTGTGGGACACACAGTGTGCTCGATGTACCTCTCGTGGGAGTTGGC 2111  
QY 2043 AACTGATGAAACAAAGAAAGTTTAAATCAGATGTTTTCCTTCCAAAGAACTCTTTGACCA 2102

Db 2112 AACTGAAAGAACAGAGAAAGTTTAAATCAGATGTTTCTCTPAAAGAAACCTTGACCA 2171  
QY 2103 TGACAGATGTGACAGCACACAGATTTGTTACAGCTGTACAGCCAAACCAATGACTGCCA 2162  
Db 2172 TGACAGATGTGACAGCACACAGATTTGTTACAGCTGTACAGCCAAATACCAATGACTGCCA 2231  
QY 2163 CTGGTGCATGACCAATGTGTGCTCCAGGAGACACAGCTGTCTCAGAGGAGCAGATCTCCAT 2222  
Db 2232 CTGGTGCATGACCAATGTGTGCTCCAGGAGACACAGCTGTCTCAGAGGAGCAGATCTCCAT 2291  
QY 2223 TTTTAGTATGACCAATTTGCCCAAGGATTAACCCCATGTCTACTGTAAACAAGAACCCAG 2282  
Db 2292 TGCCAGATGACCAATTTGCCCAAGGATTAACCCCATGTCTACTGTCAATAAGAAAACCCAG 2351  
QY 2283 CTGACAGAGCTGTGCTCCAGGAGACCAAGATCTGACAGTGGAGCCCGGAATCAGAGTGCAT 2342  
Db 2352 CTGACAGAGCTGTGCTCCAGGAGACCAAGATCTGACAGTGGAGAACCTCGAAATCAAGAGTGCAT 2411  
QY 2343 TGCCCTGCCCGAAATATCTGTGGCAATTTGGCTGGCAATTTGGTGGAACTCTCATTTTGA 2402  
Db 2412 CGCCCTGCCCGAAATATCTGTGGCAATTTGGCTGGCAATTTGGTGGAACTCTCGTGTGA 2471  
QY 2403 AATTACTACTGCAAGGAGAAATTAACAATGCTAAATTTGTTCTGTAGAAACCAATGC 2462  
Db 2472 AATCACTACTGCTAAGGAAATTAACAATGCTAAATTTGTTCTGTAGAAACCAATGC 2531  
QY 2463 CTTTTTGGCTTCTCTACACCCAGAGAGGTAGAAATTTGCTTAAACAGCTCGAAT 2522  
Db 2532 CTTTTTGGCTTCTCTACACCCAGAGAGGTAGAAATTTGCTTAAACAGCTCGAAT 2591  
QY 2523 AATGCACTCATCTCAGAGATGTCAGAGCTCACTTTAAACCCCATGGCTGGCTTGGAA 2582  
Db 2592 AATGCACTCATCTCAGAGATGTCAGAGCTCACTTTAAACCCCATGGCTGGCTTGGAA 2651  
QY 2583 GATCAATGTCTCTACTGTGGAGAGATATGTCCTTAAACCAATTTTATCAGAAACCCAGTAC 2642  
Db 2652 GATCAATGTCTCTACTGTGGAGAGATATGTCCTTAAACCAATTTTATCAGAAACCCAGTAC 2711  
QY 2643 GTGGATGCTCTGAGCCAGTGTCTGATCTGATCTGTGGAAATTTTATCAGAAACCCAGTAC 2702  
Db 2712 GTGGATGCTCTGAGCCAGTGTCTGATCTGATCTGTGGAAATTTTATCAGAAACCCAGTAC 2771  
QY 2703 TCGGAGCTGAGAGCTGTCAACCTGTCAATCAACCACTCAATGTGATGTGTGTGAAAGGCC 2762  
Db 2772 TCGGAGCTGAGAGCTGTCAACCTGTCAATCAACCACTCAATGTGATGTGTGTGAAAGGCC 2831  
QY 2763 TGCAAAACCAAGTGTGAGCAGTGTGGGACACACATGTGCTTGGAGACAGCATGTGGAGA 2822  
Db 2832 TGCAAAACCAAGTGTGAGCAGTGTGGGACACACATGTGCTTGGAGACAGCATGTGGAGA 2891  
QY 2823 TTGCACACAGCGGAGCTCTGAGTGTGATGTGGTGTGAGCAACATGAAAGCAGTGTGTGACTC 2882  
Db 2892 GTGCACACAGCGGAGCTCTGAGTGTGATGTGGTGTGAGCAACATGAAAGCAGTGTGTGACTC 2951  
QY 2883 CAATGCTATGTGGCTCTCTTCCCTTTGGCCAGTGTATGATGATGATGATGATGATGATGATG 2942  
Db 2952 CAATGCTATGTGGCTCTCTTCCCTTTGGCCAGTGTATGATGATGATGATGATGATGATGATG 3011  
QY 2943 CTGCCCCCTGAAAATTTGTTTCAAGGCTTCTGACCTGTAGTCTTGTGAGCAACCCAGG 3002  
Db 3012 CTGCCCCCTGAAAATTTGTTTCAAGGCTTCTGACCTGTAGTCTTGTGAGCAACCCAGG 3071  
QY 3003 CTGTGCTGT 3062  
Db 3072 CTGTGCTGT 3131  
QY 3063 TAAAGGACAGTGAAGATGCTTTCGAAAGCCCTTCAAGCAATTTCTATPCCACAGCCCT 3122  
Db 3132 TAAAGGACAGTGAAGATGCTTTCGAAAGCCCTTCAAGCAATTTCTATPCCACAGCCCT 3191  
QY 3123 GCTCAATTCAGAGT 3182

3192	DB	3192	TCTGAAC	TCTGATGTCTAGAGACAGCAGATACA	CTGTGCTTTTATTCACTGTCC	3251
3183	QY	3183	AGCTTCC	CAATCAACGGCCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGA	3242	
3252	DB	3252	AGCTTGC	CAGTGCATATGGACACAGCAAGTGCATCAACACAGAGCATCTGTGAGAAGTGTGA	3311	
3243	QY	3243	GAACCTG	ACACACAGGCAACGACCTGCAGAGCCTCGCATATCTGGCTTCTACGGTGATCCCAC	3302	
3312	DB	3312	GGACCTT	GCACACAGGCAAGCAGCTGCAGAGCCTCGCATATCTGGCTTCTATGGTGACCCGAC	3371	
3303	QY	3303	CAATGCA	GAGGAAATGTCAAGCATGCAATGGGCACGCGCTCTCTGTGCAACACCAA	3362	
3372	DB	3372	TAAATG	GAGGCAATGTCAAGTCGAAGTCAACGGCACGCATCACTGTGCAACACCAA	3431	
3363	QY	3363	CACGGCA	AGTCTTCTGCACACCAAGGGCGTCAAGGGGACAGATGCCAGTATGTGA	3422	
3432	DB	3432	CACGTGC	CAAGTGTCTTCTGTACCAACCAAGGCGTCAAGGGGAGAGGTGCCAGTATGTGA	3491	
3423	QY	3423	GGTAGAA	ATCGATACCAAGGAAACCTCTCAGAGGAACAATGTTATTATCTCTCTTAT	3482	
3492	DB	3492	GGTAGAA	ATCGATACCAAGGAAACCTCTCAGAGGAACAATGTTATTATCTCTCTTAT	3551	
3483	QY	3483	TGACTAT	TCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACGCTATCAA	3542	
3552	DB	3552	TGACTAT	TCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACGCTATCAA	3611	
3543	QY	3543	TTTTTGT	GGCTACTCTCGACGAAACAAACAGGGATTTGGACATGTTCAATCAATGCCCTCAA	3602	
3612	DB	3612	TTTTTGT	GGCTACTCTCGATGAAACAAACAGGGATTTGGACATGTTCAATCAATGCCCTCAA	3671	
3603	QY	3603	GAATTTCA	ACCTCAACATCACTCGGCTGCCAGTTCTCAGCTGGAAACCCAGCTCGAGA	3662	
3672	DB	3672	AAACTTT	CAACCTCAACATCACTCGGCTGCCAGTTCTCAGCTGGAAACCCAGCTCGAGA	3731	
3663	QY	3663	AGAGATG	CCCTGTTTCAAAAACCAACATTAAGGAGTACAAGAATAGTTTCTCTAATGA	3722	
3732	DB	3732	AGAGGTG	CCCTGTTTCAAAAAACCAACATTAAGGAGTACAAGAATAGTTTCTCTAATGA	3791	
3723	QY	3723	GAAGTTG	TATTTTCGCAACCAACCAATATCATCTTTCTTTGTTTATGTCAGTAATTTCA	3782	
3792	DB	3792	GAATTTG	ATTTTCGCAACCAACCAATATCATCTTTCTTTGTTTATGTCAGTAATTTCA	3851	
3783	QY	3783	CTGCCCAT	CAAAATTCA	3800	
3852	DB	3852	TTGCCCAT	CAAAATTCA	3869	

RESULT 11  
AAZ91916  
ID AAZ91916 standard; cDNA; 8827 BP.

AC AAZ91916;

08-JUN-2000 (first entry)

DE Wild type (C57BL/6J) murine mahogany protein coding sequence.

xx Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;  
 xw weight regulation; cell therapy; body weight disorder; cachexia;  
 xw anorexia; hyperpigmentation; increased metabolic rate disorder;  
 xw hyperphagia; Antibesity; antianorexic; anticachexic; ss.

SO Mus sp.

AA  
PN  
WO200005373-A2.

XX  
PD  
03-FEB-2000.

XX  
PF 21-III-1999: 99WO-US016484-

XX  
PP 21-III-1998. 98IIS-0093630P.

PR 20-OCT-1998; 98US-0104978P.  
DE 05-FEB-1999; 99US-00245041

(MILL-) MILLENIUM PHARM INC.

Moore K, Nagle DL;

WPI; 2000-195103/17.

P-PSDB; AAY81803.

New human and murine mahogany genes, useful, e.g. for diagnosis and treatment of body weight disorders.

Claim 1; Fig 2a; 188pp; English.

This sequence represents a murine mahogany gene of the invention. The mahogany genes are used: (i) to produce recombinant mahogany (mg) proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming therapeutics; (iii) as a source of diagnostic probes and primers for detecting expression of mg genes or mutations, regulatory defects, in this gene, or for isolation of related sequences; and (iv) in (cell-based) gene therapy. (ii) are used to raise specific antibodies (Ab); to identify other (extra)cellular products involved in weight regulation, and to screen for agents that disrupt interaction between (ii) and other macromolecules. The Ab are used to detect abnormal levels (or function) of (ii) for diagnosis, prognosis or monitoring of treatment; to evaluate (ii)-expressing cells intended for cell therapy, and as therapeutic mg inhibitors. Cells that express the mg gene (or contain the mg polypeptide) are used to identify agents (A) that modulate mg activity. (A) are potentially useful for the treatment of body weight disorders, particularly obesity, cachexia or anorexia, or other conditions associated with the mg gene such as hyperpigmentation, hyperphagia and disorders that result in increased metabolic rate

Sequence 8827 BP; 2304 A; 2077 C; 2162 G; 2284 T; 0 U; 0 Other;

Query Match	76.9%	Score 2938.2	DB 3	Length 8827
Best Local Similarity	85.8%	Pred. No. 0		
Matches 3399	Conservative	0	Mismatches 393	Indels 205
Gaps				

  

QY	3	GGTGGCGGACGGCGGCAACTGAGGCAAGCGCTGAGGAGGAGGACGGCGGCGACGGCAGC	62
Db	82	GGTGGCGGCGCGCGCGACTGAGGCGCGGCTGAGGGGAGGACGACGAGGACGACGAGCAGC	141

QY	63	GCTCGGGGAGAGCGCGGGCCGACNCTGGGACTGGACGTGACCAAGGCTGGAGGCC	121
DB	142	GCTTGGGGGAGAAAGGGCAGGACACCACTGCACCGGACACAGGGGCTGGAGGCC	201
QY	123	GGGCGCTGGGGGGCGGGCTGCGCCTCCCGCGCGCTGCTGTCTCCACCGCTGGGGCCACAGCT	182
DB	202	GGACACGCGCGCCCGGCTGTCTCCGCGGGTGTCTCGGGGCGCTGCCCGCCGCGCC	261
QY	183	GCTGTGCTGTGTGTGTGCTCCCGCGCGCTGTGCTGTGCTGTGCTGTGCTGTGAGGC	242
DB	262	GCTGTGCTGCTGTCTT-----TTGCTGTCTGTGCTGCTGCCGTGCCCGGGAGGC	312
QY	243	CGAGGCGCGCGCGCGCGCGGTGTCGGGCTCAGCGCAGCCAGGCGCAAGGAATG	302
DB	313	CGAGGCGCGCTGCGGTGCGCGCGCGGTGTCGGCTCGCGCGCAGCGCAGAGATG	372
QY	303	TGACCGGCGCCTGTGTCAACCGCGGTGCTGCTCAACCTGGCACCGCCAGTGCCTGCGCC	362
DB	373	TGACCGGCGCGTGTGTCAACCGGCGCGCTGCAACCTTGGCACCGCCAGTGCCTGCGCC	432
QY	363	CGCGGCTGTGGTGGGCGAGCAATGCCAGCACTGCGGGGCGCGTTCAG-----	410
DB	433	CACGGGCTGGGTGGGCGAGCAATGCCAGCACTGCGGGGCGCGTTCAGGACATGTGTCTC	492
QY	411	-----	410
DB	493	ACGCCTATAATCACAGCTGTTCCGAAAGGTGAGGCTGGAGGAACAAGTTCGAGGCAAGCTTC	552
QY	411	-----	410
DB	553	GGCTACAGAAATAAGTTCAAGAGTAACCTGGGGCAACTTGGCGTTGTCTCCAAAACCAAAA	612

QY 411 ----- 410  
Db 613 TGAGCGAAAGAGCAAGCTAGAGTCTTTGGGAAAAATTTAGCTGACTAATTTTTCACC 672  
QY 411 ----ACTAACTGGATCTCTGGTTTGTGACAGATGGACCTGGAAATTTATAAATACAAA 466  
Db 673 GAGAACTAACTGGCTCTCTCGAATTTGAACAGATGGACCTGGGAAATTTATAAATAAGA 732  
QY 467 CGAAGTGCACTGGCTCATTTGAAGGACAGCCAAATAGAAATATAGAGACTTCGTTTCAATC 526  
Db 733 CGAAGTGCACTGGCTCATTTGAAGGACAGCCAAATAGAAATATAGAGACTTCGCTTCAACC 792  
QY 527 ATTTTGTACAGAGTGTAGTTGGGACCAATATATATATATATATATATATATATATATAT 586  
Db 793 ATTTTGTACAGAACTGTAGCTGGGACCAATATATATATATATATATATATATATATATAT 852  
QY 587 CACCGCTAGTCTGCAATTTAGTGGCTCATTTGTTCCCTGAGAGATGGCAATGAGACTG 646  
Db 853 CACCTCTGATTCGCTTTAGTGGCTCATTTGTTCCCTGAAAGATGGCAATGAGACGG 912  
QY 647 TCCCTGAGTTGTTGCCAATCAGGTTATGCTTGCCTGCATTTTATAGTATGCTGCTT 706  
Db 913 CTCCTGAGTCACTGCTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 972  
QY 707 ATAAATTCAGTGGATTTAATATATCTACAGTTTGTATGATGTTCCCAATTAAGTCTCAG 766  
Db 973 ATAAATTCAGTGGATTTAATATATCTATCAATTTTGAATGTTCCGAAATTAAGTCTCAG 1032  
QY 767 GCCGAGGAGTGTAAAGATCAGTAATAGCAGCGAACTGTTGAATGTGAATGTTCTGAAA 826  
Db 1033 GCCGAGGAGTGTAAAGACAGTAACAGCAGCGCTGTTGAGTGTGAATGTTCTGAAA 1092  
QY 827 ACTGAAAGGTGAACATGTGACATTCCTCACTGTACAGAACACTGTGGTTTCTCTATC 886  
Db 1093 ACTGAAAGGGAGTGTGACATTCCTCACTGTACAGAACACTGTGGTTTCTCTCACC 1152  
QY 887 GAGGATCTGCAATTCAGATGTGACAGGATGCTCCTGCTTCTCAGACTGGCAGGGTC 946  
Db 1153 GAGGATCTGTAATCGAAGCATACAGAGGTTGCTCCTGCTTCTCAGTGGCAGGGTC 1212  
QY 947 CTGATGTTCACTCTGTACAGTAAACAGTCAATTTGACCTCGAGAGGAAATTTCTA 1006  
Db 1213 CTGATGTTCAATCTCTGTGACGTAACAGTCTTTTGGACTCGAGAGAAATTTCTG 1272  
QY 1007 ACTTAAAGCTCCAGAGCATCTCAATAAGCTGTGTCATGGAACAAATATATGTTGGTTG 1066  
Db 1273 ATTTAAAGCTTCCAGAGCTCTCAATAAGCTGTGTCATGGAATATATATGTTGGTTG 1332  
QY 1067 TTGAGGATATATGTTCAACCACTCAGATTTAATGATGTTCTAGCGTATGACCTTCTT 1126  
Db 1333 TTGCGGATATATGTTCAACCACTCAGATTTAATGATGTTCTAGCGTATGACCTTCTT 1392  
QY 1127 CTAGGAGTGGCTTCCACTAAACGTTCTGTGAACATGTTGTTTGTAGATATGTTCAAT 1186  
Db 1393 CTAGGAGTGGCTTCCACTAAACGTTCTGTGAACATGTTGTTTGTAGATATGTTCAAT 1452  
QY 1187 CTTTGGCATTTACAGGATAAAATTTACATGTATGGAGGAAAAATTTGATCCAACTGGGA 1246  
Db 1453 CTTTGGCATTTACAGGATAAAATTTACATGTATGGAGGAAAAATTTGATTTCAACAGGA 1512  
QY 1247 ATGTGACCAATGAGTTGAGATTTTCACTTCAATATGAGTCAATGGTGTGTTGACCC 1306  
Db 1513 ACGTGACCAATGAGTTGAGATTTTCACTTCAATATGATGATGATGATGATGATGATGAT 1572  
QY 1307 CTAAGGCAAGAGGAGTATGAGTGTGTTGGGCACTCTGACACATTTGTACACTGAAGA 1366  
Db 1573 CGAAGCTTAAAGATGAGTATGAGTGTGTTGGGCACTCTGACACATTTGTACACTGGAT 1632  
QY 1367 ATGGCCGAGTGTGATGCTGATCTTTTGTGCTACTGCCCTCTCTATGATATATAAGCA 1426  
Db 1633 CTGGCGTGTGATGTTGTTGCTATCTTTCGGTCACTTGGCCACTCTATGATATATAAGCG 1692

QY 1427 ATGTGAGGAAATATGATTTGGATAAGAACACATGAGGATATATATACACCCAGGGTGGCC 1486  
Db 1693 TTGTGAGGAAATATGATTTGGAAAAAACACATGAGGATATATATACATCTCAGGGTGTCT 1752  
QY 1487 TTGTGCAAGGGGTTACGGCCATAGCAGTGTGTACGACCATAGGACACGGGCTTATACG 1546  
Db 1753 TTGTGCAAGGGGTTATGGCCACAGTAGTGTGTATATGACAGGACCAAGGCTCTGTACG 1812  
QY 1547 TTCAATGCTGCTCAAAAGGCTTTTCAAGTCCAAATAGTACCGGCTTGCAGATGATCTTACC 1606  
Db 1813 TTCAATGCTGCTCAAAAGGCTTTTCAAGTCCAAATAGTACCGGCTTGCAGATGATCTTACA 1872  
QY 1607 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1666  
Db 1873 GATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1932  
QY 1667 TGCAACAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1726  
Db 1933 TGCAACAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1992  
QY 1727 ACATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1786  
Db 1993 ACATCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2052  
QY 1787 CTTGTGACCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1846  
Db 2053 CTTGTGACCGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2112  
QY 1847 GCAATTCAGCAGTCTTACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1906  
Db 2113 GCAATTCAGCAGTCTTGTACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2172  
QY 1907 TCTCTGACGACATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1966  
Db 2173 TCTCTGACGCTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2232  
QY 1967 CTTGTTTACGAGCAGGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2026  
Db 2233 CTTGTTTACGAGCAGGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2292  
QY 2027 TCTGTTGGGCTGCGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2086  
Db 2293 CTTCTGTTGGGCTGCGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2352  
QY 2087 AAAGAACTCTTGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2146  
Db 2353 AAAGAACTCTTGACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2412  
QY 2147 ACACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2206  
Db 2413 ATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2472  
QY 2207 AAGGCGCAGATCTCCTATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2266  
Db 2473 AAGGCGCAGATCTCCTATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2532  
QY 2267 GTAAACAGGACCGAGCTGAGGAGCTGCTGCTGAGCAGAACTGCGAGTGGGAGCC 2326  
Db 2533 GCAATAAGAAAAACAGCTGAGGAGCTGCTGCTGAGCAGAACTGCGAGTGGGAGCC 2592  
QY 2327 GGAATCAGGAGTGCATTCCTGCGGAAAAATATCTGTTGGCAATGCTGGGCAATTTGGTTG 2386  
Db 2593 GGAATCAGGAGTGCATTCCTGCGGAAAAATATCTGTTGGCAATGCTGGGCAATTTGGTTG 2652  
QY 2387 GAACTCTATGTTTGAATAATTTACTACTGCAAGGAGAAATATGACAAATGCTAAATGTTCT 2446  
Db 2653 GAACTCTATGTTTGAATAATTTACTACTGCAAGGAGAAATATGACAAATGCTAAATGTTCT 2712  
QY 2447 GTAGGAAACCAATGCTGCTTGTGCTTCTCTTACACCCAGGAGAGTGTAGATTTGTTCC 2506  
Db 2713 GTAGGAAACCAATGCTGCTTGTGCTTCTCTTACACCCAGGAGAGTGTAGATTTGTTCC 2772  
QY 2507 TTAAGCAGCTGCGAAATATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2566

2773 TTAAGCAGCTTCGATTAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCAT 2832  
2567 GGGTGGGCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGCTCCCATTTA 2626  
2833 GGGTTGGCTTCGGAAGATCAATGTGCTCTACTGTGCTGGGAAGATATGCTCCATTTA 2892  
2627 CAATAGATTTACTACAGTGGATGCGCTCTGAGCCAGTGTGCTGATCTCTGTGGAATTT 2686  
2893 CAATAGATTTGCTGAGTGGATGCCATCTGAGCCAGTGTGCTGATCTCTGTGGAATCT 2952  
2687 TATCAGAACCCAGTACTGGGAGCTGAAGCTGCAACCTGCACTCAACCCACTCAATGTA 2746  
2953 TGTCAAGCCTAGTACTCGGGATTTAAAGGCTGCAACCTGCACTCAACCCCTCTCAATGGCA 3012  
2747 GTGTCTGTAAAGGCTGCAACCCAGTGTAAAGCAGTGCAGGACACCATGTGCTTGA 2806  
3013 GGGTCTGTGAAGGCTGCAACCCAGTGTCCAGCAGTGCAGGACACCATGTGCTTGC 3072  
2807 GCACAGCATGTGGAATGCAACCCAGGCGAGCTCTGAGTGCATGTGGTGCAGCAACATGA 2866  
3073 GCACAGCGTGTGGGAGTGCACCTAGCAGCAGCTCGAGTGCATGTGGTGCAGTAAACATGA 3132  
2867 ASCAGTGTGACCTCCATGCTATGTGGCTCTCTCCCTTTGCGCAGTGTATGGAAT 2926  
3133 AGCAGTGTGACCTCCATGCTACGTGGCTCTCTCCCTTTGCGCAGTGTATGGAAT 3192  
2927 GGTATACGATGAGCAGCTGCCCTCCCTGAAATTTGTCAGGCTACTGTACTGTACTT 2986  
3193 GGTATACGATGAGCAGCTGCCCTCCCTGAAATTTGCTCTGGCTACTGTACTGTACCCCAT 3252  
2987 GCTTGGAGCAACAGGCTGTGCTGTGCTGTACTGATCCAGCAATACTGGCAAGGAAT 3046  
3253 GCTTGGAGCAGCAGCTGTGCTGTGCTGTACTGATCCAGCAATACTGGCAAGGAAT 3312  
3047 GCATAGAGGCTTCTTAAAGGACAGTGAAGATGCTTGGCAAGCCCTACAGGAAT 3106  
3313 GTATTGAGGCGAGCTATAAGGACCTGTGAAGTGCCTGACAGGCTCTGCAGGAATG 3372  
3107 TCTATCCAGCCCTGTCAATTCAGCATGTGCTTAGAGCAAGCAGATACAACTGGT 3166  
3373 TGTATCCAGCCCTTCTGAATCCAGCATGTGCTTAGAGCAAGCAGATACAACTGGT 3432  
3167 CTTTCACTTCACTGTCCAGCTTCCCAATGCAAGCGGCACAGTAATGCATCAATCAGACA 3226  
3433 CTTTCACTTCACTGTCCAGCTTCCCAATGCAAGCGGCACAGTAATGCATCAATCAGACA 3492  
3227 TCTGTGAGAGTGTGAGAACCTGACCAAGCAAGCAGTGCAGACCTGCAATCTGGCT 3286  
3493 TCTGTGAGAGTGTGAGAACCTGACCAAGCAAGCAGTGCAGACCTGCAATCTGGCT 3552  
3287 TCTACGGTATCCCAATGAGGGAATGTCAAGCATGTCAAGTGCATGGGACCGGT 3346  
3553 TCTATGGTACCCGACTTAATGAGGCAATGTCAAGCATGTCAAGTGCATGGGACCGGT 3612  
3347 CTTGTGCAACCAACACCGGCAAGTGTCTGCAACCAACCAAGCGGCTCAAGGGGAGC 3406  
3613 CACTGTGCAACCAACACCGGCAAGTGTCTGCAACCAACCAAGGCTCAAGGGGAGC 3672  
3407 AGTGCCAGTATGTAGGTAGAAATCGATACAAGGAACCCCTCTCAGAGGAACATGTT 3466  
3673 AGTGCCAGTATGTAGGTAGAAATCGATACAAGGAACCCCTCTCAGAGGAACATGTT 3732  
3467 ATTATACTCTCTTATTGACTATCACTTCACTTGTCTATCCAGGAAGATGATCGT 3526  
3733 ACTATACCCCTTCTCATTTGACTATCACTTCACTTGTCTATCCAGGAAGATGATCGT 3792  
3527 ATTACAGCATATCAATTTTGTGGCTACTCTGACGAACAAACAGGGATTTGGACATGT 3586  
3793 ACTACAGCATATCAATTTTGTGGCTACTCTGATGAACAAACAGGGATTTGGACATGT 3852  
3587 TCAATCAATGCTCCAGAAATTCACCTCAACATCACTGGGCTGCCAGTTCTCAGCTG 3646

Db 3853 TCATCATGCTCCAAAACTTCAACCTCAACATCACTGGGCCACCAAGCTTCCCAGCG 3912  
Qy 3647 GAACCCAGGCTGGAGAGAGATGCTGTTGTTTCAAAAACCAACATTAAAGAGTACAAAG 3706  
Db 3913 GAACCCAGACTGGAGAGAGGTCGCTGTTTCAAAAACCAACATCAAGATACAAAG 3972  
Qy 3707 ATAGTTTCTCTAATGAGAACTTTGATTTTGCACACCCCAAAATATACATCTTTTGT 3766  
Db 3973 ATAGTTTCTCTAATGAGAACTTTGATTTTGCACACCCCAAAATATACATCTTTTGT 4032  
Qy 3767 ATGTCAGTAATTTTCACTGGCCCATCAAAATTCAGT 3803  
Db 4033 ATGTCAGTAATTTTCACTGGCCCATCAAAATTCAGT 4069

## RESULT 12

AAS72660

ID AAS72660 standard; cDNA; 3490 BP.

XX AAS72660;

XX 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #8464.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US9008631.

XX 31-MAR-2000; 2000US-00540217.

XX 23-AUG-2000; 2000US-00649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

XX P-PSDB; ABG08473.

XX New isolated polynucleotide and encoded polypeptides, useful in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits and to assess

XX biodiversity.

XX Claim 1; SEQ ID NO 8464; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

XX sequences. (I) is useful as hybridisation probes, polymerase chain

XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,

XX and in recombinant production of (II). The polynucleotides are also used

XX in diagnostics as expressed sequence tags for identifying expressed

XX genes. (I) is useful in gene therapy techniques to restore normal

XX activity of (II) or to treat disease states involving (II). (II) is

XX useful for generating antibodies against it, detecting or quantitating a

XX polypeptide in tissue, as molecular weight markers and as a food

XX supplement. (II) and its binding partners are useful in medical imaging

XX of sites expressing (II). (I) and (II) are useful for treating disorders

XX involving aberrant protein expression or biological activity. The

XX polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic

XX coding sequences of the invention. Note: The sequence data for this

XX patent did not appear in the printed specification, but was obtained in

XX electronic format directly from WIPO at







1954 TGTTAGAACCAACATGCGCTTTTGGCTTCTCTTACAAACCAGAGAGAGTGGAAATTTGT 2013  
2505 CCTTAAG - CAGCTGCGAATAATGCA - GTCATCTCAGA - GCAATGTCAG - GCTCACTTAA 2559  
2014 CCTTAAGCGAGGTGCGAATAATGCGAGTCACTCAGAGCATGTCCAGGCTCACTTAA 2073  
2560 ACCCATATGG - TCGGCTTCGGAAGATCAA - TGTGTCTACTCTGCTGCTGGAAGATAT 2615  
2074 ACCCATATGGGCTCGGCGCTTCGGAAGATCAAATTTGTCTCTACTCTGCTGCTGGAAGATAT 2133  
2616 GTCCCCATTTACAAATAGTTTACTACAGTGTGATGCGCTCTGAGCCAGTATGCTGGATT 2675  
2134 GTCCCCATTTACAAATAGTTTACTACAGTGTGATGCGCTCTGAGCCAGTATGCTGGATT 2193  
2676 GTTGGGAATTTT - ATCAGAACCCAGTACTCGGGAGCTGAAGGCTGCAACTGTCAT - 2729  
2194 CTGTGGGAATTTTATTTCAGGAACCCAGTTACTTCGGGAGCTGAAGGCTGCAACTGTCAT 2253  
2730 -CAACCCACTCAATGCTAGTGTCTGTGA - GSCCTGCAAAACACAGT - GCTAAGCAGT 2785  
2254 CAACCCACTCAATGCTAGTGTCTGTGAAGGCTGCAAAACACAGTGTGTTAAGCAGT 2313  
2786 GCCGACACCATGTGCC - TTAGGACACAGTATGAGATTTGACACAGCGGAGCTCT 2841  
2314 GCCGACACCATGTGCCCTTTGAGGGACAGCATGTGGAGATTGACACAGCGGAGCTCT 2373  
2842 GAGT - GCAATGCTGAGCAACATGAA - GCAATGCTGAGTCTCAATGCTATGTCGCT 2899  
2374 GAGTGGCATGTGTGTCAGCAACATGAAGGCACTGTGTGACTCCAAATGCCATATGTCGCT 2433  
2900 CTTTCCCTTTT - GGCAGTGTATGCAATGGTATACAGTGA - GCACCTGCCCTCCCTGAAAA 2957  
2434 CTTTCCCTTTTGGCCAGTGTATGGAATGATACAGTATGAGCACTGCCCTGAAAA 2493  
2958 TTGTTACGGTACTGTACCTGTAGTCAATGCTGAGCAACCA - GGCTGTGGCTGTGTTA 3016  
2494 TTGTTACGGTACTGTACCTGTAGTCAATGCTGAGCAACCAAGGCTGTGGCTGTGTTA 2553  
3017 CTGATCCAGCAATCTGGCAAGGGAATCATAGAGGTTCTATAAAGGACCACTGA 3076  
2554 CTGATCCAGCAATCTGGCAAGGGAATCATAGAGGTTCTATAAAGGACCACTGA 2613  
3077 AGATGCTTTCGCAAGCCCTCAGGAAATTTCTATCCACAGCCCTGCTCAATTTCCAGCA 3136  
2614 AGATGCTTTCGCAAGCCCTCAGGAAATTTCTATCCACAGCCCTGCTCAATTTCCAGCA 2673  
3137 TGTGTCTAGAGCAGCAGATACACTGG - TCTTTCACTTCACTGTCCAGCTTCCCAATG 3194  
2674 TGTGTCTAGAGCAGCAGATACACTGGTCTTTTCATTTCACTGTCCAGTTTCCCAATG 2733  
3195 CAAC - GGCACACAGTAAATGCAATCAGAGCATCTGTGAGAAGTGTGAGAACCCTGACCA 3253  
2734 GAAAGGCAACAGTAAATGCAATCAGAGCATCTGTGAGAAGTGTGAGAACCCTGACCA 2793  
3254 CAGGCAAGCAGTCCGAGACCTGCATATCTGGCTTCTAC - GGTGATCCACCAA - TGGAGG 3311  
2794 CAGGCAAGCAGTCCGAGACCTGCATATCTGGCTTCTACGGGTGATCCCAATTTGGAGG 2853  
3312 GAAATGTGAGCAGTCAATGCGCAAGGCGCTCTGTGCAACACCAACACAGGCGCA 3371  
2854 GAAATGTGAGCAGTCAATGCGCAAGGCGCTCTGTGCAACACCAACACAGGCGCA 2913  
3372 GTGCTTCTGACCAACCAAGGCGCTCAAGGGGACGAGTGCCAGTATGTAGGTTAGAAAA 3431  
2914 GTGCTTCTGACCAACCAAGGCGCTCAAGGGGACGAGTGCCAGTATGTAGGTTAGAAAA 2973  
3432 TCGATACCAAGAAACCTCTCAGAGGAACATGTTATTATATCTCTCTTATTGACTATCA 3491  
2974 TCGATACCAAGAAACCTCTCAGAGGAACATGTTATTATATCTCTCTTATTGACTATCA 3033  
3492 GTTCACTTTTGTCTATCCAGGAAGATGATGCTATTACAGCTATCAATTTTGTGGC 3551  
3034 GTTCACTTTTGTCTATCCAGGAAGATGATGCTATTACAGCTATCAATTTTGTGGC 3093

3552 TACTCTGACGACAAACACAGGATTTGACATGTTTCATCATGCTCCAGAAATTTCAA 3611  
3094 TACTCTGACGACAAACACAGGATTTGACATGTTTCATCATGCTCCAGAAATTTCAA 3153  
3612 CTTCAACATCACTCGGCTGCGAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCC 3671  
3154 CTTCAACATCACTCGGCTGCGAGTTTCTCAGCTGGAACCCAGGCTGGAGAAGAGATGCC 3213  
3672 TCTTGTTCCTCAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTATGAGAAGTTTGA 3731  
3214 TCTTGTTCCTCAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTATGAGAAGTTTGA 3273  
3732 TTTTTCGCAACCAACCAATATCACTTTCTTTTATGTCAGTAATTTTCACTGCGCCCAT 3791  
3274 TTTTTCGCAACCAACCAATATCACTTTCTTTTATGTCAGTAATTTTCACTGCGCCCAT 3333  
3792 CAAAATTCAGGTGCA 3806  
3334 CAAAATTCAGGTGAA 3348  
RESULT 13  
AAZ91922  
ID AAZ91922 standard; cDNA; 2625 BP.  
XX  
AC AAZ91922;  
XX  
XX 08-JUN-2000 (first entry)  
XX  
XX Human mahogany protein coding sequence #4.  
DE  
XX  
XX Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;  
KW weight regulation; cell therapy; body weight disorder; cachexia;  
KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200005373-A2.  
PN  
XX  
PD 03-FEB-2000.  
XX  
XX 21-JUL-1999; 99WO-US016484.  
PF  
XX  
XX 21-JUL-1998; 98US-0093630P.  
PR  
XX  
XX 20-OCT-1998; 98US-0104978P.  
PR  
XX  
XX 05-FEB-1999; 99US-00245041.  
XX  
XX (MILL-) MILLENIUM PHARM INC.  
PA  
XX  
XX Moore K, Nagle DL;  
PI  
XX  
XX WPI; 2000-195103/17.  
XX  
XX P-PSDB; AAY81809.  
DR  
XX  
XX New human and murine mahogany genes, useful, e.g. for diagnosis and  
PT treatment of body weight disorders.  
PS  
XX  
XX Claim 1; Fig 20a; 188pp; English.  
XX  
XX This sequence represents a human mahogany gene of the invention. The  
CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
CC proteins (II); (iii) as a source of antisense, ribozyme or triplex-forming  
CC therapeutics; (iii) as a source of diagnostic probes and primers for  
CC detecting expression of mg genes or mutations, regulatory defects, in  
CC this gene, or for isolation of related sequences; and (iv) in (cell-  
CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to  
CC identify other (extra)cellular products involved in weight regulation,  
CC and to screen for agents that disrupt interaction between (II) and other  
CC macromolecules. The Ab are used to detect abnormal levels (or function)  
CC of (II) (for diagnosis, prognosis or monitoring of treatment); to  
CC evaluate (II)-expressing cells intended for cell therapy, and as



Db 1980 ACCTGGTATTCGGTGTGTGTGGAAACACAGGGTCTGCTCAGTGTATCTCGTGGCGCTGGC 2039  
 Qy 2043 AACTGATGAACAGAGAAAGTTAAATCAGAAATGTTTCCAAAGAACTCTTGACCA 2102  
 Db 2040 AACTGATGAACAGAGAAAGTTAAATCAGAAATGTTTCCAAAGAACTCTTGACCA 2099  
 Qy 2103 TGACAGATGTGACACAGACAGATGTTTACAGCTGTACAGCCAAACCAATGACTGGCA 2162  
 Db 2100 TGACAGATGTGACACAGACAGATGTTTACAGCTGTACAGCCAAACCAATGACTGGCA 2159  
 Qy 2163 CTGCTGCAATGACCAATGTTTCCAGAGAACACAGCTGCTGAGAGGCCAGATCTGCAT 2222  
 Db 2160 CTGCTGCAATGACCAATGTTTCCAGAGAACACAGCTGCTGAGAGGCCAGATCTGCAT 2219  
 Qy 2223 TTTTAGGTATCAGAAATGCCCCCAAGGATAACCCATGTAATCTACTGTAAACAAGACACAG 2282  
 Db 2220 TTTTAGGTATGAGAAATGCCCCCAAGGATAACCCATGTAATCTACTGTAAACAAGACACAG 2279  
 Qy 2283 CTGACAGAGCTGTGCTTGGACACAGAACTGCCAGTGGAGAGCCCGGGAATCAGGAGTGCAT 2342  
 Db 2280 CTGACAGAGCTGTGCTTGGACACAGAACTGCCAGTGGAGAGCCCGGGAATCAGGAGTGCAT 2339  
 Qy 2343 TGGCCTGCCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGG 2387  
 Db 2340 TGGCCTGCCCGTAGGCTTGCAGGGTCTATCTTGGTGTGTGGG 2384

## RESULT 14

AAZ91917  
 ID AAZ91917 standard; cDNA; 2419 BP.

AC AAZ91917;

XX 08-JUN-2000 (first entry)

XX Murine mahogany protein coding sequence akml003.

DE Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;  
 XX weight regulation; cell therapy; body weight disorder; cachexia;  
 KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
 KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.

XX Mus sp.

XX WO200005373-A2.

XX 03-FEB-2000.

XX 21-JUL-1999; 99WO-US016484.

XX 21-JUL-1998; 98US-0093630P.

PR 20-OCT-1998; 98US-0104978P.

PR 05-FEB-1999; 99US-00245041.

XX (MILL-) MILLENIUM PHARM INC.

XX Moore K, Nagle DL;

XX WPI; 2000-195103/17.

DR P-PSDB; AAY81804.

XX New human and murine mahogany genes, useful, e.g. for diagnosis and

PT treatment of body weight disorders.

XX Claim 1; Fig 8a; 188pp; English.

XX This sequence represents a murine mahogany gene of the invention. The  
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iii) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in cell-

CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (ii) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (ii) for diagnosis, prognosis or monitoring of treatment); to  
 CC evaluate (ii)-expressing cells intended for cell therapy, and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other  
 CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate  
 XX

SQ Sequence 2419 BP; 578 A; 567 C; 680 G; 594 T; 0 U; 0 Other;

Query Match 39.8%; Score 1521.8; DB 3; Length 2419;

Best Local Similarity 80.7%; Pred. No. 0;

Matches 1899; Conservative 0; Mismatches 247; Indels 205; Gaps 2;

Qy 3 GGTGGCGGACGCGCGGCAACTGAGCAAGGCTGAGGAGGAGCGGCGGCAACGGCAGC 62

Db 82 GGTGGCGGCGCGCGGCAACTGAGCGCGGCTGAGGGGGAGCAGAGACGACAGCAGC 141

Qy 63 GCTCGCGGCGAGGAGCGGCGCGGCACTGGGACTGGGACGTCGACAGGGCTGGGAGGCC 122

Db 142 GCTCGCGGCGAGGAGCGGCGCGGCACTGGGACTGGGACGTCGACAGGGCTGGGAGGCC 201

Qy 123 GGGGCTGGGGCGCGGCTCGGCTCCGCGGCTGTGTCTCCACCGCTCGGCGCCACGGCT 182

Db 202 GGGACCGCGCGCGGCTGTGTCTCCGCGGCTGTGTCTCGCGGCGCTGCCCGCCGCC 261

Qy 183 GCTGCTGCTGTGTGTGTCTCCGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 242

Db 262 GCTGCTGCTGT 312

Qy 243 CGAGGCGCGGCGGCGGCGGCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 302

Db 313 CGAGGCGCGGCTGT 372

Qy 303 TGACCGCGGCTGT 362

Db 373 TGACCGCGGCTGT 432

Qy 363 CGCGGCTGT 410

Db 433 CACGGCTGT 492

Qy 411 ----- 410

Db 493 ACGCTATATACAGCTGTTCGGAAGTGTAGGCTGGAGGAACAGTTCGAGCAAGCTTC 552

Qy 411 ----- 410

Db 553 GGCTACAGATAAGTTCAAGATAACCTGGGCGCACTTGGGCTTGTCTCCAAACCAAAA 612

Qy 411 ----- 410

Db 613 TGAGCGAAAGAGGCAAGCTAGAGTCTTTTGGGAAAAATTTAGCTGAGTAATTTTTCACC 672

Qy 411 ----ACTAACTGGATCTTCTGGGTTTGTACAGATGGACCTGGAAATTAATAACAAAA 466

Db 673 GAGAACTACTGGCTCTTCTGGATTGTAAACAGATGGACCTGGGAATTAATAATAGA 732

Qy 467 CGAGTGTACGTGGCTCATTGAAGGACAGCCAAATAGAAATATGAGCTTGTTCATTC 526

Db 733 CGAAGTGCATGGCTCATTTGAAGGACAGCCAAATAGAAATATGAGCTTGTTCATTC 792

Qy 527 ATTTTCTACAGATGTAGTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATG 586

Db 793 ATTTTCTACAGATGTAGTGGGACCAATTTATATGTTTATGATGGGACTCAATTTACG 852

Qy 587 CACCGCTAGTCTGTGATTTAGTGGCTCATTTGTTCCTGAGAGATGCGAATGAGACTG 646





Sat Mar 6 08:28:40 2004

us-09-787-097-19.rng

Page 32

Job time : 917.913 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	3713.8	97.2	4072	9	US-09-893-	238-16	Sequence 16, Appl
2	3696.2	96.8	8589	9	US-09-893-	238-14	Sequence 14, Appl
3	2939.8	77.0	8827	9	US-09-893-	238-1	Sequence 1, Appl
4	2248.2	58.9	2625	9	US-09-893-	238-18	Sequence 18, Appl
5	1521.8	39.8	6319	9	US-09-893-	238-8	Sequence 8, Appl
6	1364.4	35.7	6370	9	US-09-893-	238-12	Sequence 12, Appl
7	1525	22.1	6733	16	US-10-197-	824-1	Sequence 1, Appl
8	640.2	16.8	1051	9	US-09-893-	238-10	Sequence 10, Appl
9	277.4	7.3	625	14	US-10-138-	845-5869	Sequence 5869, Ap
10	245.2	6.5	531	9	US-09-884-	761-15900	Sequence 15900, A
11	246	6.4	246	9	US-09-736-	457-966	Sequence 966, App
12	246	6.4	246	9	US-09-902-	941-966	Sequence 966, App
13	246	6.4	246	9	US-09-849-	626-966	Sequence 966, App
14	246	6.4	246	14	US-10-017-	754-966	Sequence 966, App
15	246	6.4	246	14	US-10-113-	878-966	Sequence 966, App

09 67 78 78 78

DD  
09





Db 2280 CTGCAGAGTGTGCTCCCTGACGACGAACTGCCAGTGGAGCCCGGAAATCAGAGTGCAT 2339  
Qy 2343 TGCCTCCCGGAAATATCTGTGGCAATTTGGCTGCGCATTTGGTTGGAAACTCATGTTGAA 2402  
Db 2340 TGCCTCCCGGAAATATCTGTGGCAATTTGGCTGCGCATTTGGTTGGAAACTCATGTTGAA 2399  
Qy 2403 AATTACTACTGCCAAGAGAAATATGACAAATGCTAAATTTCTGTAGGAAACCAATGC 2462  
Db 2400 AATTACTACTGCCAAGAGAAATATGACAAATGCTAAATTTCTGTAGGAAACCAATGC 2459  
Qy 2463 CCTTTGGCTTCTCTTACAAACCCAGAAAGAGTGAATTTGTCTTAAGCAGCTGGCAAT 2522  
Db 2460 CCTTTGGCTTCTCTTACAAACCCAGAAAGAGTGAATTTGTCTTAAGCAGCTGGCAAT 2519  
Qy 2523 AATCAGTCTACTCAGAGCATGTCCTCAAGCTCACTTAACCCCATGGGTGGCTTCGGAA 2582  
Db 2520 AATCAGTCTACTCAGAGCATGTCCTCAAGCTCACTTAACCCCATGGGTGGCTTCGGAA 2579  
Qy 2583 GATCAATGTGCTTACTGCTGCGAAGATATGCTCCCATTTTACAAATAGTTTACTACA 2642  
Db 2580 GATCAATGTGCTTACTGCTGCGAAGATATGCTCCCATTTTACAAATAGTTTACTACA 2639  
Qy 2643 GTGATGCCCTCTCAGAGCATGCTGCTGGAATCTGTGGAATTTTATCAGAACCCATAC 2702  
Db 2640 GTGATGCCCTCTCAGAGCATGCTGCTGGAATCTGTGGAATTTTATCAGAACCCATAC 2699  
Qy 2703 TCGGGACTGAAGCTGCAACTGCATCAACCCACTCAATGGTAGTCTGTGAAAGGCC 2762  
Db 2700 TCGGGACTGAAGCTGCAACTGCATCAACCCACTCAATGGTAGTCTGTGAAAGGCC 2759  
Qy 2763 TGCAAACACAGTGTAGCAGTGTCCGACACCACTATGCTGCTTGAGGACAGCATGTGAGA 2822  
Db 2760 TGCAAACACAGTGTAGCAGTGTCCGACACCACTATGCTGCTTGAGGACAGCATGTGAGA 2819  
Qy 2823 TTGCACAGCGGAGCTCTGAGTGTGATGTGAGCAACATGAAGCAGTGTGGAATC 2882  
Db 2820 TTGCACAGCGGAGCTCTGAGTGTGATGTGAGCAACATGAAGCAGTGTGGAATC 2879  
Qy 2883 CAATGCCCTATGGCTCTCTTCCCTTTGGCCAGTGTATGAAATGCTATCAGATGAGCAC 2942  
Db 2880 CAATGCCCTATGGCTCTCTTCCCTTTGGCCAGTGTATGAAATGCTATCAGATGAGCAC 2939  
Qy 2943 CTGCCCCCTGAAATTTGTCAGGCTACTGTACTGTAGTCACTGTGAGTGTGAGCAACAGG 3002  
Db 2940 CTGCCCCCTGAAATTTGTCAGGCTACTGTACTGTAGTCACTGTGAGTGTGAGCAACAGG 2999  
Qy 3003 CTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGAAATGCTATGAGGTTCTTA 3062  
Db 3000 CTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGAAATGCTATGAGGTTCTTA 3059  
Qy 3063 TAAAGGACCAAGTGAAGATGCTTCCGAAAGCCCTTACAGGAAATTTCTATCCAGCCCT 3122  
Db 3060 TAAAGGACCAAGTGAAGATGCTTCCGAAAGCCCTTACAGGAAATTTCTATCCAGCCCT 3119  
Qy 3123 GCTCAATTCAGATGTGTAGAGGACAGCATGATCAATCAGATGCTGTGCTTCTTACTGTCC 3182  
Db 3120 GCTCAATTCAGATGTGTAGAGGACAGCATGATCAATCAGATGCTGTGCTTCTTACTGTCC 3179  
Qy 3183 AGCTTGCCAAATCAAGGCAAGTAAATGCTATCAGATGCTGTGAGAGTGTGA 3242  
Db 3180 AGCTTGCCAAATCAAGGCAAGTAAATGCTATCAGATGCTGTGAGAGTGTGA 3239  
Qy 3243 GAACCTGACCAAGGCAAGCACTGCGAGACCTGCAATATCTGGCTTCTACGGTGTATCCAC 3302  
Db 3240 GAACCTGACCAAGGCAAGCACTGCGAGACCTGCAATATCTGGCTTCTACGGTGTATCCAC 3299  
Qy 3303 CAATGAGGAAATGTCAGGCAAGTAAATGCTATGAGGACAGCTGTGTCACACCA 3362  
Db 3300 CAATGAGGAAATGTCAGGCAAGTAAATGCTATGAGGACAGCTGTGTCACACCA 3359  
Qy 3363 CACGGCAAGTGTCTGTCACCAACCAAGGCGCTCAAGGGGACGAGTGTGCTATGTGA 3422  
Db 3360 CACGGCAAGTGTCTGTCACCAACCAAGGCGCTCAAGGGGACGAGTGTGCTATGTGA 3419

Qy 3423 GGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTATCTCTCTTAT 3482  
Db 3420 GGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTATCTCTCTTAT 3479  
Qy 3483 TGAATATCAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAA 3542  
Db 3480 TGAATATCAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAA 3539  
Qy 3543 TTTTGTGGCTACTCTCTGAGCAACAAACAGGAAATTTGGACATGTTTCATCAATGCTTCCAA 3602  
Db 3540 TTTTGTGGCTACTCTCTGAGCAACAAACAGGAAATTTGGACATGTTTCATCAATGCTTCCAA 3599  
Qy 3603 GAATTTCAACCTCAACATCACTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3662  
Db 3600 GAATTTCAACCTCAACATCACTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3659  
Qy 3663 AGAGATGCTGTGTTTCAAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTTAATGA 3722  
Db 3660 AGAGATGCTGTGTTTCAAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTTAATGA 3719  
Qy 3723 GAAGTTTGAATTTTCGCAACCAACCAATATCACTTCTTGTGTTATGTCAGTAAATTCAC 3782  
Db 3720 GAAGTTTGAATTTTCGCAACCAACCAATATCACTTCTTGTGTTATGTCAGTAAATTCAC 3779  
Qy 3783 CTGGCCCATCAAAATTCAGTGTGCAACTGAAACAATGA 3819  
Db 3780 CTGGCCCATCAAAATTCAGTGTGCAACTGAAACAATGA 3816

## RESULT 2

US-09-893-238-14  
; Sequence 14, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Regie, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 8589  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-893-238-14

Query Match 96.8%; Score 3696.2; DB 9; Length 8589;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 3744; Conservative 0; Mismatches 48; Indels 9; Gaps 1;  
Qy 3 GGTGGCCCGCAGCGCGGCAACTCAGGCAAGGCTGAGGAGGAGGACGCGCGGCGACGCGC 62  
Db 9 GGTGGCCCGCAGCGCGGCGCAACTCAGGCAAGGCTGAGGAGGAGGACGCGCGGCGACGCGC 68  
Qy 63 GCTCGCGGCGAGGAGCGCGCGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 122  
Db 69 GCTCGCGGCGAGGAGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 128  
Qy 123 GGGGCTCGGGGCGCGGCGGCTGCGCGCGGCGGCTGCTCTCCACCGCTGCGGCGGCGGCGG 182  
Db 129 GGGACCGCGCGCGGCGGCTGCTCTCCCGCGGCGGCTGCTCGCGGCGGCGGCTGCGCGGCGG 188  
Qy 183 GCTGCTGCTGCTGTTGTTGCTCCCGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242



QY	2403	AAATTACTCTGCCAAGGAGAAATTATGACAATGCTAAATTTGTTCTGTGAGGAACCAATGC	2462
Db	2400	AAATTACTCTGCCAAGGAGAAATTATGACAATGCTAAATTTGTTCTGTGAGGAACCAATGC	2459
QY	2463	CCTTTTGGCTTCTCTTCAACCCAGAAGAAGTAGAAATTTTGCTTTAAACAGCTCGGAAT	2522
Db	2460	CCTTTTGGCTTCTCTTCAACCCAGAAGAAGTAGAAATTTTGCTTTAAACAGCTCGGAAT	2519
QY	2523	AATGCAGTCACTCTCAGAGCATGTCGAAGCTCACTTAAACCCATGGCTCGGCTTCGGAA	2582
Db	2520	AATGCAGTCACTCTCAGAGCATGTCGAAGCTCACTTAAACCCATGGGTGGCTTCGGAA	2579
QY	2583	GATCAATGTCTCTTACTGTGTGCTGGGAAGATATGTCCCATTTTACAAATAGTTTACTACA	2642
Db	2580	GATCAATGTCTCTTACTGTGTGCTGGGAAGATATGTCCCATTTTACAAATAGTTTACTACA	2639
QY	2643	GTGGATGCCGTCTGAGCCGAGTGATGCTGGATTTCTGTGGNAATTTATCAGAACCCAGTAC	2702
Db	2640	GTGGATGCCGTCTGAGCCGAGTGATGCTGGATTTCTGTGGNAATTTATCAGAACCCAGTAC	2699
QY	2703	TCGGGGACTGAAGGCTCAACCTGCACTCAACCCACTCAATGTGTAGTGTCTGTGAAAGGCC	2762
Db	2700	TCGGGGACTGAAGGCTCAACCTGCACTCAACCCACTCAATGTGTAGTGTCTGTGAAAGGCC	2759
QY	2763	TGCAAAACACAGTGTCTAAGCAGTGC CGGAACCAATGTGCTTGGAGCAGCATGTGGAGA	2822
Db	2760	TGCAAAACACAGTGTCTAAGCAGTGC CGGAACCAATGTGCTTGGAGCAGCATGTGGAGA	2819
QY	2823	TTGCAACACGGGAGCTCTGAGTGCAATGCTGGTCAGCAACATCAAGCAGTGTGTGACATC	2882
Db	2820	TTGCAACACGGGAGCTCTGAGTGCAATGCTGGTCAGCAACATCAAGCAGTGTGTGACATC	2879
QY	2883	CAATGCCCTATGTGGCCCTCCITTCCTTTGGCCAGTGTATGGAATGGTATACGATAGCAC	2942
Db	2880	CAATGCCCTATGTGGCCCTCCITTCCTTTGGCCAGTGTATGGAATGGTATACGATAGCAC	2939
QY	2943	CTGCCCCCTTGAAAATTTGTTTTCAGGCTACTGTACCTGTAGTCAATTCCTTGAGCAACACAGG	3002
Db	2940	CTGCCCCCTTGAAAATTTGTTTTCAGGCTACTGTACCTGTAGTCAATTCCTTGAGCAACACAGG	2999
QY	3003	CTGTGGCTGGTGTACTGATCCAGCAATCTGGCAAGGGAATGCAATAGAGGTTTCCTTA	3062
Db	3000	CTGTGGCTGGTGTACTGATCCAGCAATCTGGCAAGGGAATGCAATAGAGGTTTCCTTA	3059
QY	3063	TAAAGGACAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGGCCCT	3122
Db	3060	TAAAGGACAGTGAAGATGCTTTCGCAAGCCCTTACAGGAAATTTCTATCCACAGGCCCT	3119
QY	3123	GCTCAATTCAGACATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCAATTCACATGTCC	3182
Db	3120	GCTCAATTCAGACATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCAATTCACATGTCC	3179
QY	3183	AGCTTGCCAATGCAACGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTGA	3242
Db	3180	AGCTTGCCAATGCAACGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTGA	3239
QY	3243	GAACCTGACACAGCAAGCACTGCGAGACCTGCAATCTGGCTTCTACGGTGTATCCAC	3302
Db	3240	GAACCTGACACAGCAAGCACTGCGAGACCTGCAATCTGGCTTCTACGGTGTATCCAC	3299
QY	3303	CAATGGAGGGAATGTGAGCCATGCAAGTGAATGGGCACGCTCTCTGTGCAACACCAA	3362
Db	3300	CAATGGAGGGAATGTGAGCCATGCAAGTGAATGGGCACGCTCTCTGTGCAACACCAA	3359
QY	3363	CACGGGCAAGTCTTCTGCAACCAACGGGGCTCAAGGGGACAGGTGCCAGCTATGTGA	3422
Db	3360	CACGGGCAAGTCTTCTGCAACCAACGGGGCTCAAGGGGACAGGTGCCAGCTATGTGA	3419
QY	3423	GGTGAATAATGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTTATCTCTCTTAT	3482
Db	3420	GGTGAATAATGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTTATCTCTCTTAT	3479

Qy	3483	TGACTATCAGTTCCACCTTTAGTCTATCCCGGAAGATGATCGGTATTATACACAGCTATCAA	3542
Db	3480	TGACTATCAGTTCCACCTTTAGTCTATCCCGGAAGATGATCGGTATTATACACAGCTATCAA	3539
Qy	3543	TTTTTGGCTACTCCCTGACGAAACAAACACGGGATTTGGACATGTTTCATCAATGCCTCCAA	3602
Db	3540	TTTTTGGCTACTCCCTGACGAAACAAACACGGGATTTGGACATGTTTCATCAATGCCTCCAA	3599
Qy	3603	GAATTTCAACCTCAACATCACCCTGGGCTGCCAGTTTCTCAGCTGGGAACCCAGCGCTGAGA	3662
Db	3600	GAATTTCAACCTCAACATCACCCTGGGCTGCCAGTTTCTCAGCTGGGAACCCAGCGCTGAGA	3659
Qy	3663	AGAGATGCCTGTTGTTTCAAAAACCAACATTAAAGAGGTACAAAGATGATTTCTCTAATGA	3722
Db	3660	AGAGATGCCTGTTGTTTCAAAAACCAACATTAAAGAGGTACAAAGATGATTTCTCTAATGA	3719
Qy	3723	GAGATTGATTTCGGCAACCAACCAATATACATTTCTTTGTTTATGTCAGTAATTTTCAC	3782
Db	3720	GAGATTGATTTCGGCAACCAACCAATATACATTTCTTTGTTTATGTCAGTAATTTTCAC	3779
Qy	3783	CTGCCCATCAAAATTCAGGT	3803
Db	3780	CTGCCCATCAAAATTCAGAT	3800

### RESULT 3

```

US-09-893-238-1
; Sequence 1, Application US/09893238
; Patent No. US2020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS
; TITLE OF INVENTION: DIAGNOSIS OF BODY WE
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSQ for Windows Version 3.0.0
; SEQ ID NO 1
; LENGTH: 8827
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-1

```

Query Match	77.0%;	Score 2939.8;	DB 9;	Length 8827;
Best Local Similarity	85.1%;	Pred. No. 0;		
Matches 3400;	Conservative	0;	Mismatches 392;	Indels 205; Gaps 12
Qy	3	GCTGGCCGACGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGCGACGGCAGC	62	
Db	82			
Qy	63	GCTCGCGGCAGGACGGCGGGCCGCACTGGGACTGGGACGTGACACAGGGCTGGGAGGCC	122	
Db	142	GCTTGGGGCAGGAAGGCGACGACGCCACCCCTGCACCGCAGCAGGGGGCTGGAGGCC	201	
Qy	123	GGGGCTGGGGGCGGGCTGCGCCTCCCGCGGCTGCTGTCTCCACCGCTTGGGCCACGGCT	182	
Db	202	GGACCGCGCCCGGCTGTGTCTCCCGGGGTGTGTTCGCGGGCGCTGTCCCGCCCGCCGCC	261	
Qy	193	GCTGCTGCTGTGTGTGTGCTCCCGCGCGCTGTGTGCTGTGCTGTGCTGTGTGAGGC	242	
Db	262	GCTGTGCGCGTGTCTT-----TTGCTGCTGTGTGCTGCGCTGTGCCCGGAGGC	312	
Qy	243	CGAGGCCGCGGGCGGGCGGGGTTCGCGGTCAGCCCGCAGCGGCGCAAGGAATG	302	

Db 313 CGAGGCGCTGTGGTGGCGCGCGGTGTCCGGCTCGGCGGACGCGAGGCCAAGGAATG 372  
QY 303 TGACCGGCGCTGTGTCAACGGCGGTGTGTGACACCTGTGCGACCGCGCGGTGTGTGCCC 362  
Db 373 TGACCGGCGGTGTGTCAACGGCGCGGTGTGTGACACCTGTGCGACCGCGCGGTGTGTGCCC 432  
QY 363 CGCGGCTGGTGGGCGAGCAATGCGAGCACTGCGGGGCGCGCTTCAG----- 410  
Db 433 CACGGCTGGTGGGCGAGCAATGCGAGCACTGCGGGGCGCGCTTCAGGACATCTGTCTC 492  
QY 411 ----- 410  
Db 493 AGCCCTAATCACAGCTGTTCCGAAGGTGAGGCTGGAGAAACAGTTCCAGGCAAGCTTC 552  
QY 411 ----- 410  
Db 553 GGCACAGAAATAGTTCAAGAGTAACCTGGGGCACTGGGGCTGTCTCCAAACCAAAA 612  
QY 411 ----- 410  
Db 613 TGAGCGAAAGAGCAAGCTAGAGTCTTTTGGGAAAAATTTAGCTGCACTAATTTTTCACC 672  
QY 411 -----ACTAACTGGATCTCTCGGTTTGTGACAGATGGACCTGGAAATTAATAACAAAA 466  
Db 673 GAGAACTAACTGGCTCTCTGGATTTGTAAACAGATGGACCTGGGAAATTAATAATAAGA 732  
QY 467 CGAAGTGCACTGGCTCATTTGAAGGACAGCCAAATAGAAATATGAGACTTCGTTTCAATC 526  
Db 733 CGAAGTGCACTGGCTCATTTGAAGGACAGCCAAATAGAAATATGAGACTTCGTTTCAACC 792  
QY 527 ATTTTGTACAGAGTGTAGTGGGACCAATTAATATATGTTTATGATGGGCACTCAATTTATG 586  
Db 793 ATTTTGTACAGAGTGTAGTGGGACCAATTAATATGTTTATGATGGGCACTCAATTTATG 852  
QY 587 CACCGCTAGTCTGCATTTAGTGGCTCATTTGTTCTCGAGAGATGGCAATGAGACTG 646  
Db 853 CACCTCTGATTTGCTGCTTTAGTGGCTCATTTGTTCTCGAAGAGATGGCAATGAGACGG 912  
QY 647 TCCCTGAGTTGTTCCACATCAGTTATGCTGCTGCTCATTTTATGATGATGCTGCTT 706  
Db 913 CTCCTGAGTCACTGCTCACTTCAGGTTATGCACTGCTGCATTTTTCAGTATGCTGCTT 972  
QY 707 ATAAATGACTGGATTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 766  
Db 973 ATAAATGACTGGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1032  
QY 767 GCCGAGGAGTGTAAAGTACAGTAAATAGCAGCGAAACGTTGTAATGTAATGTTCTGAAA 826  
Db 1033 GCCGAGGAGTGTAAAGTACAGTAAATAGCAGCGCGCTGTTGAGTGTGAATGTTCTGAAA 1092  
QY 827 ACTGGAAGGTGAAAGTACAGTAAATAGCAGCGCGCTGTTGAGTGTGAATGTTCTGAAA 886  
Db 1093 ACTGGAAGGTGAAAGTACAGTAAATAGCAGCGCGCTGTTGAGTGTGAATGTTCTGAAA 1152  
QY 887 GAGGATCTGCAATTCAGTAAATAGTACAGGATGCTCTGCTTCTCAGACTGGCAGGGTC 946  
Db 1153 GAGGATCTGCAATTCAGTAAATAGTACAGGATGCTCTGCTTCTCAGACTGGCAGGGTC 1212  
QY 947 CTGATGTTCTGCTGTACAGTAAATAGTACAGGATGCTCTGCTTCTCAGACTGGCAGGGTC 1006  
Db 1213 CTGATGTTCTGCTGTACAGTAAATAGTACAGGATGCTCTGCTTCTCAGACTGGCAGGGTC 1272  
QY 1007 ACTTAAAGCTCCAGAGCATCTCATAAAGCTGTGGTCAATGGAACATTTATGCGGTTG 1066  
Db 1273 ATTTAAAGCTCCAGAGCATCTCATAAAGCTGTGGTCAATGGAATATTAATGCGGTTG 1332  
QY 1067 TTGAGGATATATGTTCAACACTCAGATTAATAAGTGTCTGAGCTGTGAGCTTGTGCTT 1126  
Db 1333 TTGCGGATATATGTTCAACACTCAGATTAATAAGTGTCTGAGCTGTGAGCTTGTGCTT 1392  
QY 1127 CTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACCAATGTTGTTTATGATATGTTGCTT 1186  
Db 1393 CTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACCAATGTTGTTTATGATATGTTGCTT 1452

QY 1187 CTTTGGCATTTATACAGGATAAATTTTACATGTATGGAGGAAAAATTTGATCCAATGGGA 1246  
Db 1453 CTTTGGCATTTATACATAGGATAAATTTTACATGTATGGAGGAAAAATTTGATCCAAGGGA 1512  
QY 1247 ATGTGACCAATGAGTTGAGAGTTTTCACATTCATATAGTCAATGGTGTGTTGTTGACCC 1306  
Db 1513 ACGTGACCAATGAGTTGAGAGTTTTCATATTCATATGAATCATGGTATTTGTTAACTC 1572  
QY 1307 CTAAGGCAAAAGGAGCAGTATGCAAGTGTGGGCACTCTGCACACATTTGTACTCAAGA 1366  
Db 1573 CGAAAGTAAAGGATCAGTATGCAAGTGTGGGCACTCAGCACACATTTGTACTGCGAT 1632  
QY 1367 ATGCCGAGTGGTCACTGCTGCTCATCTTTGGTCACTGCCCCCTCTATGATATATAAGA 1426  
Db 1633 CTGGCGTGTGTCATGTTGGTCACTCTTGGTCACTTGGCCACTCTATGATATATAAGG 1692  
QY 1427 ATGTGACGGAATATGATTTGGATPAAGAACACATGAGATATATTACACACCCAGGTTGCC 1486  
Db 1693 TTGTGACGGAATATGATTTGGAAAGAACACATGAGATATATTACATCTCAGGTTGCTC 1752  
QY 1487 TTGTGCAAGGGGTTAGGCCATAGCAGTGTGTTTACACCATAGGACGAGGCCCTATAG 1546  
Db 1753 TTGTGCAAGGGGTTATGGCCACAGTAGTGTGTTTATGATGACAGGACCAAGGCTCTGTACG 1812  
QY 1547 TTCATGGTGTCTACAAAGGCTTTTCAGTCCCAATAAGTACCAGGCTTGAGATGATCTCTACC 1606  
Db 1813 TTCATGGTGTCTACAAAGGCTTTTCAGGCGCAACAAATACCGCTTGAGATGACCTCTACA 1872  
QY 1607 GATATGATGTGATACCCAGATGTGACCATTCATTTAAGACAGCCGATTTTTCGTTACT 1666  
Db 1873 GATACATGTGATATCTCAGATGTGACCATTCATTTAAGACAGCCGATTTTTCGTTACT 1932  
QY 1667 TGACACAGCTGTGATGAGTGAAGAACCATGCTGTGTTTGGGGGAAACACACACAATG 1726  
Db 1933 TGACATACAGCTGTGATGAGTGAAGAACCATGCTGTGTTTGGAGGAAACACACACAATG 1992  
QY 1727 ACACATCTATGAGCCATGCGCCAAATGCTTCTTTAGATTTTATGGCTTATGACATG 1786  
Db 1993 ACACCTTCCATGAGCCACGGTGCACAAATGCTTCTCGGCTTCATGGCTTATGACATG 2052  
QY 1787 CCTGTGACCGCTGTGTCAGTGTCCAGACCTGATCTCCACCATGATGTCACACATTTG 1846  
Db 2053 CTTGTGACCGCTGTGTCAGTGTCCAGACCTGATCTCCACCATGATGTCACACATTTG 2112  
QY 1847 GCATTTACAGCTGTGTTACACACAGCACCATGATGTTTGGTGTGTTTCAATAGTCTCC 1906  
Db 2113 GCATTTACAGCTGTGTTACACACAGCACCATGATGTTTGGTGTGTTTCAATAGTCTCC 2172  
QY 1907 TCCACGACCATCTGTTGATTTACCTCGGACAGTGTGATGCGCATGGAGTGAAGCG 1966  
Db 2173 TCTCAGTACGCTTGTGTTTACCTCGGACAGTGTGATGCAACCGCAGTGAAGCTG 2232  
QY 1967 CTTGTTTACAGCAGGACCTGTTATTTGCTGTGTGGAACACAGGCTGCTCTCAGTGT 2026  
Db 2233 CTTGTTGTCAGCAGGACCTGTTATTTCCGTTGCTGTGGGACACACAGTCTCTCGATGTA 2292  
QY 2027 TCTGTTGGGCGCTGCAACTGATGAAACAGAAAGTTAAATCAAGATGTTTTTCCA 2086  
Db 2293 CTTCTCGGAGTTGGCAACTGAAAGAACAGCAAGAAAGTTAAATCAGATGTTTTTCTA 2352  
QY 2087 AAAGAACTCTTACCATCAGATGTGACACAGCACACAGATTTGTACAGCTGTACAGCCA 2146  
Db 2353 AAAGAACTCTTACCATCAGATGTGACACAGCACACAGATTTGTACAGCTGTACAGCCA 2412  
QY 2147 ACACCAATGACTGCCACTGCTGTCATGACCATGCTGTCCTCCAGGAAACACAGCTGCTCAG 2206  
Db 2413 ATACCAATGACTGCCACTGCTGTCATGACCATGCTGTCCTCCAGGAAACACAGCTGCTCAG 2472  
QY 2207 AAGGCCAGATCTCCATTTTATGATGAGAAATGCCCCCAAGGATAACCCCATGTACTACT 2266  
Db 2473 AAGGCCAGATCTCCATTTTATGATGAGAAATGCCCCCAAGGATAACCCCATGTACTACT 2532

2267 GTACAGAGACCGAGCTGCGAGAGCTGTCGCCCTGGACCAAGATGCGAGTGGAGGCC 2326  
QY  
2533 GCAATAAGAAACACAGCTGCGAGAGCTGTCGCCCTAGACCAAGATGCGAGTGGAGGCC 2592  
Db  
2327 GGAATCAGAGTGCATTCGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTG 2386  
QY  
2593 GGAATCAGAGTGCATTCGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTG 2652  
Db  
2387 GAAATCTATGTTGAAATTAATCTACTGCCAAGGAGAAATATGACAATGCTAAATGTTCT 2446  
QY  
2653 GAAATCTGCTGTAATTAATCACTACTGTAAGGAGAAATATGACAATGCTAAATGTTCT 2712  
Db  
2447 GTAGAAACCAATGCCCTTTTGGCTTCTTACAAACCCAGAGAGAGTAAATTTGTCC 2506  
QY  
2713 GTAGAAACCAATGCCCTTTTGGCTTCTTACAAACCCAGAGAGAGTAAATTTGTCC 2772  
Db  
2507 TTAAGCAGCTGCAATTAATGCAATCTCAGAGCATGTCCAAGCTCACTTAAACCCCAT 2566  
QY  
2773 TTAAGCAGCTGCAATTAATGCAATCTCAGAGCATGTCCAAGCTCACTTAAACCCCAT 2632  
Db  
2567 GGGTCGGCTTCGGAAGATCAATGTGCTTACTGTGCTGGAGAGATGTCCCATTTA 2626  
QY  
2833 GGGTTGGCTTCGGAAGATCAATGTGCTTACTGTGCTGGAGAGATGTCTCCATTC 2892  
Db  
2627 CAATAGTTTACTACAGTGGATGCCGCTGTGAGCCAGTGTGCTGATTTCTGTGGAATTT 2686  
QY  
2893 CAATAGTTTGTGCAATGGATGCCATCTGAGCCAGTGTGCTGCTTCTGTGGATCT 2952  
Db  
2687 TATCAGAACCCAGTACTCGGAGATGAAGGCTGCAACCTGTCAACCCATCAATGGTA 2746  
QY  
2953 TGTCAAGCGCTAGTACTCGGAGTAAAGGCTGCAACCTGTCAACCCATCAATGGCA 3012  
Db  
2747 GTGCTGTGAAGGCTTCRAACCACTGCTTAAGCTGCGGACCACTGTCCTTGA 2806  
QY  
3013 GCGTCTGTGAAGGCTTCRAACCACTGCTTAAGCTGCGGACCACTGTCCTTGC 3072  
Db  
2807 GGACAGCATGTGGAGATTCACACGCGCAGCTCTGAGTGTGCTGTCGACCAATGA 2866  
QY  
3073 GGACAGCATGTGGAGATTCACACGCGCAGCTCTGAGTGTGCTGTCGACCAATGA 3132  
Db  
2867 AGCAGTGTGAGCTTCAATGCTATGTCGCTTCTTCCCTTTGGCCAGTGTATGGAT 2926  
QY  
3133 AGCAGTGTGAGCTTCAATGCTATGTCGCTTCTTCCCTTTGGCCAGTGTATGGAT 3192  
Db  
2927 GGTATACGATGAGCACTGCCCTGCTGAAATTTGTCAGGCTACTGTCACCTGTAGTCAAT 2986  
QY  
3193 GGTATACGATGAGCACTGCCCTGCTGAAATTTGTCAGGCTACTGTCACCTGTAGTCAAT 3252  
Db  
2987 GCTTGAGCAACCGGCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3046  
QY  
3253 GCTTGAGCAACCGGCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3312  
Db  
3047 GCATAGAGGTTCTTAAAGGACCAAGTGAAGATGCTTCCGACCCCTACAGGAATTT 3106  
QY  
3313 GTATTAGGGAGCTATAAGGACCTGTGAAGATGCTTCCGACCCCTACAGGAATTT 3372  
Db  
3107 TCTATCCAGCGCCCTGCTCAATTCAGCATGTGCTAGAGGACAGCAGATACACTGGT 3166  
QY  
3373 TGTATCCAGCGCCCTTCTGAACTCCAGCATGTGCTAGAGGACAGCAGATACACTGGT 3432  
Db  
3167 CTTTCAATTCAGTTCAGCTTCCCAATGCAAGCGGACCAAGTAAATGCAATGCAATGCA 3226  
QY  
3433 CTTTCAATTCAGTTCAGCTTCCCAATGCAAGCGGACCAAGTAAATGCAATGCAATGCA 3492  
Db  
3227 TCTGTGAGAAAGTGTGAGAACTTACCACAGGCAAGCACTGCGAGACCTGTCATATCTGGCT 3286  
QY  
3493 TCTGTGAGAAAGTGTGAGAACTTACCACAGGCAAGCACTGCGAGACCTGTCATATCTGGCT 3552  
Db  
3287 TCTACGGTATCCCAATGAGGGAATGTGAGCAATGCAAGTGAATGCAATGAGGCAAGCT 3346  
QY  
3553 TCTATGGTGAACCGCAATGAGGCAATGTGAGCAATGCAAGTGAATGCAATGAGGCAAGCT 3612  
Db  
3347 CTTCTGTGCAACCAACAGCGGCAAGTGTCTTCTGCAACCAACAGGCGCTCAAGGGGACG 3406  
QY

3613 CACTGTGCAACACCAACACCGCAAGTGTCTTGTACCAACCAAGTGTCAAGGGGACG 3672  
Db  
3407 AGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTT 3466  
QY  
3673 AGTGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGCT 3732  
Db  
3467 ATTATATCTTCTTATGACTATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCT 3526  
QY  
3733 ACTATACCTTCTCATGACTATCAGTTCACTTTAGCTGTCTCCAGGAAGACGCGCT 3792  
Db  
3527 ATTACACAGCTATCAATTTTGTGGCTACTCTGACGAAACAAACACAGGATTTGGACATGT 3586  
QY  
3793 ACTACACAGCTATCACTTTTGTGGCTACTCTGATGAAACAAACACAGGATTTGGAATGT 3852  
Db  
3587 TCATCAATGCTCCAAAGATTTCAACCTCAACATCACTTGGGCTGCGAGTTTCTCAGCTG 3646  
QY  
3853 TCATCAATGCTCCAAAGATTTCAACCTCAACATCACTTGGGCTGCGAGTTTCTCAGCTG 3912  
Db  
3647 GAACCCAGCTGAGAGAGATGCTGTTTCAAAAACCAACATTAAGGATACAAAG 3706  
QY  
3913 GAACCCAGCTGAGAGAGATGCTGTTTCAAAAACCAACATTAAGGATACAAAG 3972  
Db  
3707 ATAGTTTCTCAATGAGAGTTTGAATTTGGCAACCAACCAATATCACTTTCTTTGTTT 3766  
QY  
3973 ATAGTTTCTCAATGAGAGTTTGAATTTGGCAACCAACCAATATCACTTTCTTTGTTT 4032  
Db  
3767 ATGTCAGTAATTTCACTGCGCCATCAAAATTCAGT 3803  
QY  
4033 ATGTCAGTAATTTCACTGCGCCATCAAAATTCAGAT 4069  
Db

## RESULT 4

US-09-893-238-18  
; Sequence 18, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: Fast-Seq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 2625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-893-238-18

Query Match 58.9%; Score 2248.2; DB 9; Length 2625;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 2308; Conservative 0; Mismatches 68; Indels 9; Gaps 1;  
QY 3 GGTGGCGCAGCGCGGCAACTGAGGCAAGGCTGAGAGAGAGACGCGCGGCGAGCGGACG 62  
Db 9 GGTGGCGCAGCGCGGCAACTGAGGCAAGGCTGAGAGAGAGACGCGCGGCGAGCGGACG 68  
QY 63 GCTCGCGGCGAGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 122  
Db 69 GCTCGCGGCGAGGAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 128  
QY 123 GGGGCTGGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 182  
Db 129 GGGACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 188













Db 1926 CAAACACTTGGCTGTTCAGAAAATAAAGAGAGCTATTGTACAAAGGTGATATAGCCATA 1985  
QY 1511 GCAGTGTTTACGACCATAGACAGAGCCCTATACGTTTCATGGTGGCTACAGGCTTTCA 1570  
Db 1986 CTAGTGTGATGATGAATAACAAAGTCCATTTATGTTTCATGAGGGTATAAGCAATTCG 2045  
QY 1571 GTGCCAATAAGTACCGGCTTGACAGATCTCTACCGATATGATGTGATACCCAGATGT 1630  
Db 2046 CAGGGAAACAAATATGGAATGGTGTGATGATCTTTATAAATATGAAGTTAACTAAGACT 2105  
QY 1631 GGAACATCTTAAGGACACCGGATTTTCCGTTACTTGCACACAGCTGTGATGTAGTGTG 1690  
Db 2106 GGAATTTTGAAGAAAGTGGTTGGCAGATACCTTCATTCAGCTGTTCTTATCAATG 2165  
QY 1691 GAACCATGCTGGTGTGGGGGAAACACACAAATGACACATCTATAGAGCAATGGCGCA 1750  
Db 2166 GAGCTATGCTTATTTTGGAGGAAATACCCATAATGACACTTCTCTTGAGTAACGGTGCA 2225  
QY 1751 AATGCTTCTTTCAGATTTTCATGGGCTATGACATTCCTGTGACCGCTGTGAGTCTTC 1810  
Db 2226 AATGTTTTCTCCGATTTCTGGCATATGACATAGCTTGTGATGAATGGAATACTAC 2285  
QY 1811 CAGACCTGATCTCCACCATATGTCACACAGATTTGGCCATTCAGCAGCTTTTACACACA 1870  
Db 2286 CAAAACCAATCTTCATAGAGATGTCAACAGATTTGGACACTCTGACAGTAGTCATTAACG 2345  
QY 1871 GCACCATGTATGTTGCGTGGTTTCAATAGTCTCTCTCAGCAGACATCTCTGATNTCA 1930  
Db 2346 GTCCATGATATATTTTGGGGATTTCTAGTGTACTCTTAATGATATCTTGTATACA 2405  
QY 1931 COTCGAAACAGTGTGATGCGCATCGAGTGAAGCGCTGTTTGTAGCAGCAGACTGTGTA 1990  
Db 2406 AGCTCCAAATGCAAGGCTTTCACAGATGAAGACTTTGTAAATACTGCTGTCAGGGA 2465  
QY 1991 TTCGTGTGTGGAAACACAGGCTGCTCTAGTGTATCTGTTGGGCGCTGCAACTGATG 2050  
Db 2466 TAAATGTGTTTGGAAATAAATCACTGTGAAT-----CTTGGGAATCTGGGA 2513  
QY 2051 AACAGAGAAAGTTAAATCAGAAATGTTTTCAAAAGAACTCTTGACCATCAGACAGAT 2110  
Db 2514 ATACTAAATATTTCTAGAGCAAGTCCCTCTTAAACAGCTGCTTCTGATGACAGAT 2573  
QY 2111 GTGACGACACACAGATTTTACAGCTGTACAGCAACACCAATGACTGCCACTGGTGA 2170  
Db 2574 GTTACAGATATGCAATGTGTCAGCTGTACTGTGCAATACAAATGGGTGCCAATGGTGTG 2633  
QY 2171 ATGACCATGTGTCCTCCAGGAACACAGCTGCTCAGAGGGCCAGATCTCCATTTTAGGT 2230  
Db 2634 ATGACAGAAATGCAATTCGGCAATATGTAATCTGAGTATGTCTGTCAGAACTACACCA 2693  
QY 2231 ATGAGAAATGGCCCAAGGATAACCCATGTACTACTGTAAACAGACAGCTGCGAGGA 2290  
Db 2694 AATGTCATGTGAGAAATGA-----GCAGATTTGTAAACAACTTACAGCTGTAAA 2744  
QY 2291 GCTGTGCCCTGACACAGACTGCGAGTGGAGCCCGGAATCAGAGTGCATTTGCCCTGC 2350  
Db 2745 GCTGTTCACTAACTTGAATGCGCAGTGGATCAGACAGCAAGATGCCAGCTTTAC 2804  
QY 2351 CCGAAATATCTGTGGCAATGGCTGGCAATTTGGTTGGAACTCATGTTTGAATAATTA 2410  
Db 2805 CAGCTCATCTTTGTGGAGAGATGGAGTCAATTTGGGATGCTTGTCTTAGACTCAAT 2864  
QY 2411 CTGCAAGAGAAATATGACAAATGTAAATTTGTTCTGTAGGAACCAACAATGCCCTTTTGG 2470  
Db 2865 CCAGTAGAGAAATATGACAAATGCAAACTTTATGCTATAATCTTTAGTGGAAATCTTG 2924  
QY 2471 CTCTCTTACACCCAGAGAGAGTGTAGAAATTTGCTCTTAAGCAGCTGCGAATAATGCA 2530  
Db 2925 CTTGATTAACAACTTCAAAAGAGTAGAATTTGTT-----TGATG 2966  
QY 2531 CATCTCAGAGCATGTCCAAAGTCACTTAACTTAACTTAACTTAACTTAACTTAACTTAA 2590  
Db 2967 AAATACAGAGATATACACACAGAAAGTATCACCTTGGTAGGTTCGGCAAGATCAATA 3026

QY 2591 TGTCTACTGTGCTGGGAAGATATGTCCCATTTTAAATAGTTTTACTACAGTGATGC 2650  
Db 3027 TATCTATTGGGATGGGAAGACATGTCTCTTTTACAAACACACACTACAGTGCTTC 3086  
QY 2651 CGTCTGAGCCAGTGTGCTGGATTTCTGTGAATTTTATCAGAACCCAGTACTCGGGGAC 2710  
Db 3087 CTGCGGAACCAATGATTTCTGGGTTTTTGTGCATATCTGGAAGGGCTGAGTGGCAGGT 3146  
QY 2711 TGAAGGTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCTG----- 2764  
Db 3147 TAAAGCTAATCTTGTACATCTATGCAAAATGCGCTTGTCTGTGAAAACCTGTTGTTA 3206  
QY 2765 ---CAAACACAGTGTAGAGAGTGGGACACCATGTGCTGAGCAAGCATGTGGAG 2821  
Db 3207 GTCCAAATCAAAATGGAGCGCTGCAAAAGCATGCTCTCTGAGGACATCATGTTC 3266  
QY 2822 ATTGCACAGCGGCTCTGAGTGTGCTGAGCAACATGAAGCAGTGTGTGACT 2881  
Db 3267 ACTGTACAGCAATGCGATGGAGTGTATGTGTGAGCAGTACGAAACGATGTTGACT 3326  
QY 2882 CCAATGCTATGTGGCTCTCTCTTTTGGCAGTGTATGGAATGTTATACGATGACA 2941  
Db 3327 CTAATGCTATATCATCTCTTTTCCATATGGAATGTCTAGATGGCAAC---TGCCA 3383  
QY 2942 CCTGCCCTCTGAATTTGTTTCCAGGCTACTGTACCTGTGTAGTCTATGCTTGGAGCAAC 3001  
Db 3384 CTTGCTCTCTCTTAAATTTGTTCTGGAATGGAACCTGTGAGCAGTGTGGAACAGCTG 3443  
QY 3002 GCTGTGGCTGTGACTCCAGCAATCTGCAAGAACTGCAAGGAAATGCTAGAGGTTCT 3061  
Db 3444 GATGTGCTGTGCAATGATCTTAGTAAATACAGGAAGAGACATGCAATGGAAGTTCT 3503  
QY 3062 ATAAAGCAGTGAAGATGCTTCCAGAGCCCTACAGGAATTTCTATCCACAGCCCC 3121  
Db 3504 CACGGGACCAATGAAG-----CTTATTGGAATGCCAACATGATGG 3548  
QY 3122 TGTCTAATTCAGCATGTGTGTAGAGCAGCAGATACAACTGGTCTTTTCAATCTCATGTC 3181  
Db 3549 TTTCTGACCAATCTTGTGCCCAAGAAAGAACTATAGTGTGCTCTTATCCAGTGTG 3608  
QY 3182 CAGCTTGCAATGCAAGGCGCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTG 3241  
Db 3609 CAGCTTGCCAGTGAATGACATAGCACTTGCATCAATAATATGTGTGCGAACAGTGT 3668  
QY 3242 AGAACTGACACAGCAAGCACTGAGACCTGATATCTGCTTCTACGCTGATCCCA 3301  
Db 3669 AATCTCTACCAAGAAAGCAGTGTCAAGTTGTATGCCAGTTATTTATGAGATCCAA 3728  
QY 3302 CCAATGAGGGAATGTGAGCCATCAAGTGCATGGCAATGGGCAAGCTCTCTGTGCAACCA 3361  
Db 3729 CCAATGGTGACAGTGCACAGCTTGTACATGCAAGTGGCCATGCAAAATATCTCATCTGC 3788  
QY 3362 ACAGGGCAAGTCTTCTGCACCAACAGGCGCTCAGGGGAGCAGTGGCAGCTATGTG 3421  
Db 3789 ACACAGAAAAATGTTTCTGCAAACTAAAGGAATAAAGGTGACCAATGCGCAATATGTG 3488  
QY 3422 AGGTAGAAATCGATACCAAGAAACCTCTCAGAGGAACATGTTTATATCTCTTTA 3481  
Db 3849 ACTCTGAAATCGCTATGTTGTAATCCACTTATAGGAACATGTTTATACAGCTTTGA 3908  
QY 3482 TTGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTTATTTACAGCATCA 3541  
Db 3909 TTGATATCAATTTTCACTTTTACAGTTTATTTACAGGAAGATGATGCCACCATACTGCCATA 3968  
QY 3542 ATTTTGTGCTACTCTCTGACGAAACAAACAGGGAATTTGCAATGTTTCAATGCTTCCA 3601  
Db 3969 ACTTTATAGCAACCCAGCAACAGTGCACAAAATCTGGATATATCAATTTATGATCAA 4028  
QY 3602 AGAATTTCAACTCAACATCACTCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGAG 3661  
Db 4029 ACAACTTAACTCAACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4088



Db 220 GAGATTGCACCGGCGAGCTCTGAGTGTATGTGGTGGAGCAACATGAAGCAGTGTGG 161  
QY 2879 ACTCCAAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGA 2938  
Db 160 ACTCCAAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGA 101  
QY 2939 GCACCTGCCCCCTGAAAATTTCTCAGGCTACTGTACCT 2977  
Db 100 GCACCTGCCCCCTGAAAATTTCTCAGGCTACTGTACCT 62

## RESULT 10

US-09-864-761-15900  
; Sequence 15900, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aomic-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 15900  
; LENGTH: 531  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL132773.14  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9

US-09-864-761-15900  
; Sequence 15900, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aomic-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 15900  
; LENGTH: 531  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL132773.14  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9

Query Match 6.5%; Score 249.2; DB 9; Length 531;  
Best Local Similarity 96.9%; Pred. No. 9.6e-62;  
Matches 254; Conservative 0; Mismatches 8; Indels 0; Gaps 0;  
QY 1195 TTATACAGGATATAAATTTACATGTATGGAGGAAAAATTCATCCAACTGGGAATGTGACC 1254  
Db 141 TTTTCTAGGATATAAATTTACATGTATGGAGGAAAAATTCATCCAACTGGGAATGTGACC 200  
QY 1255 AATGAGTTGAGAGTCTTTTTCACATTCATATGAGTCTCATGGTGTGTGACCCCTAAGGCA 1314  
Db 201 AATGAGTTGAGAGTCTTTTTCACATTCATATGAGTCTCATGGTGTGTGACCCCTAAGGCA 260  
QY 1315 AAGGAGCAGTATGCACTGTGGGCACTCTGCACACATTTTACACTGAAGAATGCGCA 1374  
Db 261 AAGGAGCAGTATGCACTGTGGGCACTCTGCACACATTTTACACTGAAGAATGCGCA 320  
QY 1375 GTGCTCATGCTGCTCATCTTTTGGTCACTGCGCTCTCTATGGATATATAGCAATGTGCA 1434  
Db 321 GTGCTCATGCTGCTCATCTTTTGGTCACTGCGCTCTCTATGGATATATAGCAATGTGCA 380  
QY 1435 GAATATGATTGGATAAGACA 1456  
Db 381 GAATATGATTGGTAGGTATA 402  
RESULT 11  
US-09-736-457-966/c  
; Sequence 966, Application US/09736457  
; Patent No. US20020168637A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Iodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; APPLICANT: Wang, Aijun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C15  
; CURRENT APPLICATION NUMBER: US/09/736,457  
; CURRENT FILING DATE: 2000-12-13  
; NUMBER OF SEQ ID NOS: 1864  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 966  
; LENGTH: 246  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-736-457-966

Query Match 6.4%; Score 246; DB 9; Length 246;  
Best Local Similarity 100.0%; Pred. No. 4.6e-61;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2515 CTGCGAATAATGCACTCATCTCAGAGCATGCCAAGCTCACCTTAAACCCATGGTGGC 2574  
Db 246 CTGCGAATAATGCACTCATCTCAGAGCATGCCAAGCTCACCTTAAACCCATGGTGGC 187  
QY 2575 CTTGGAAGATCAATGTGTCTTCTACTGTGTGGGAAGATATGTCCTCCATTTTACAAATAGT 2634  
Db 186 CTTGGAAGATCAATGTGTCTTCTACTGTGTGGGAAGATATGTCCTCCATTTTACAAATAGT 127  
QY 2635 TTACTACAGTGGATGCGGTCTGAGCCAGTATCTGGAATCTGGAATTTATCAGAA 2694  
Db 126 TTACTACAGTGGATGCGGTCTGAGCCAGTATGCTGGATTTGTGGAATTTTATCAGAA 67  
QY 2695 CCCAGTACTCGGGGACTGAAGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGT 2754  
Db 66 CCCAGTACTCGGGGACTGAAGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGT 7



```
QY 2575 CTTGGAAGATCAATGTGTCCTTACTGGTCTGGGAAGATATGTCCCAATTTACAATAAGT 2634
Db |||||
186 CTTGGAAGATCAATGTGTCCTTACTGGTCTGGGAAGATATGTCCCAATTTACAATAAGT 127
QY 2635 TTACTACAGTGGATGCGCTGAGCCCAAGTGTGATGCTGGATCTGTGGAAATTTATCAGAA 2694
Db |||||
126 TTACTACAGTGGATGCGCTGAGCCCAAGTGTGATGCTGGATCTGTGGAAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAGGCTCAACCTCCATCAACCCCACTCAATGGTAGTGTCTGT 2754
Db |||||
66 CCCAGTACTCGGGGACTGAAGGCTCAACCTCCATCAACCCCACTCAATGGTAGTGTCTGT 7
QY 2755 GAAAGG 2760
Db |||||
6 GAAAGG 1
```

## RESULT 15

```
US-10-113-872-966/c
; Sequence 966, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-113-872-966
```

```
Query Match 6.4%; Score 246; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGTGGC 2574
Db |||||
246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGTGGC 187
QY 2575 CTTGGAAGATCAATGTGTCCTTACTGGTCTGGGAAGATATGTCCCAATTTACAATAAGT 2634
Db |||||
186 CTTGGAAGATCAATGTGTCCTTACTGGTCTGGGAAGATATGTCCCAATTTACAATAAGT 127
QY 2635 TTACTACAGTGGATGCGCTGAGCCCAAGTGTGATGCTGGATCTGTGGAAATTTATCAGAA 2694
Db |||||
126 TTACTACAGTGGATGCGCTGAGCCCAAGTGTGATGCTGGATCTGTGGAAATTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAGGCTCAACCTCCATCAACCCCACTCAATGGTAGTGTCTGT 2754
Db |||||
66 CCCAGTACTCGGGGACTGAAGGCTCAACCTCCATCAACCCCACTCAATGGTAGTGTCTGT 7
QY 2755 GAAAGG 2760
Db |||||
6 GAAAGG 1
```

Search completed: March 5, 2004, 13:18:59  
Job time : 857.236 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	Time				
1	3713.8	97.2	4072	3	US-09-245-041-16		Sequence 16, Appl
2	3696.2	96.8	8589	3	US-09-245-041-14		Sequence 14, Appl
3	2939.8	77.0	8827	3	US-09-245-041-1		Sequence 1, Appl
4	2248.2	58.9	2625	3	US-09-245-041-18		Sequence 18, Appl
5	1521.8	39.8	2419	3	US-09-245-041-8		Sequence 8, Appl
6	1364.4	35.7	6370	3	US-09-245-041-12		Sequence 12, Appl
7	640.2	16.8	1051	3	US-09-245-041-10		Sequence 10, Appl
8	246	6.4	246	4	US-09-702-705-966		Sequence 966, App
9	246	6.4	246	4	US-09-702-705-966		Sequence 966, App
10	246	6.4	246	4	US-09-614-124B-966		Sequence 966, App
11	246	6.4	246	4	US-09-671-325-966		Sequence 966, App
12	212.8	5.6	415	4	US-09-833-381-1916		Sequence 1916, Ap
13	201.2	5.3	90050	3	US-09-245-041-5		Sequence 5, Appl
14	169.2	4.4	17056	3	US-09-245-041-3		Sequence 3, Appl
15	132.2	3.5	226	3	US-09-245-041-6		Sequence 6, Appl
16	115.4	3.0	221	3	US-09-245-041-7		Sequence 7, Appl
17	100.2	2.6	5973	3	US-09-245-041-4		Sequence 4, Appl
18	72.8	1.9	477	3	US-09-135-984-1		Sequence 1, Appl
19	72.8	1.9	477	4	US-09-684-843A-1		Sequence 1, Appl
20	72.6	1.9	10348	2	US-08-457-273B-41		Sequence 41, Appl
21	72.6	1.9	10348	3	US-08-556-419-13		Sequence 13, Appl
22	72.6	1.9	10348	3	US-09-041-886-14		Sequence 14, Appl
23	72.6	1.9	10366	1	US-08-246-982A-5		Sequence 5, Appl
24	72.6	1.9	10366	1	US-08-453-285-5		Sequence 5, Appl
25	70.4	1.8	6530	2	US-08-146-930-1		Sequence 1, Appl
26	70.4	1.8	6530	3	US-08-458-240-1		Sequence 1, Appl
27	70.4	1.8	6530	5	PCF-US93-03993-1		Sequence 1, Appl





QY 2523 AATGACATCTCAGAGCATGTCCTAAGCTCACCTTAACCCCATGGTGGGCTTCGGAA 2582  
Db 2520 AATCAGTCTATCTCAGAGCATGTCCTAAGCTCACCTTAACCCCATGGTGGGCTTCGGAA 2579  
QY 2583 GATCAATGTGTCTTACTGCTGCTGGAGATATGTCCCAATTTACAAATAGTTTACTACA 2642  
Db 2580 GATCAATGTGTCTTACTGCTGCTGGAGATATGTCCCAATTTACAAATAGTTTACTACA 2639  
QY 2643 GTGATGCGCTCTGAGCGCAGTGTGCTGCTGGAGATATGTCCCAATTTACAAATAGTTTACTACA 2702  
Db 2640 GTGATGCGCTCTGAGCGCAGTGTGCTGCTGGAGATATGTCCCAATTTACAAATAGTTTACTACA 2699  
QY 2703 TCGGGGACTGAGGCTGCAACTGCACTCAACCCACTCAATGGTAGTGTCTGTGAAGGCC 2762  
Db 2700 TCGGGGACTGAGGCTGCAACTGCACTCAACCCACTCAATGGTAGTGTCTGTGAAGGCC 2759  
QY 2763 TGCACCAAGTGTCTAAGAGTGGCGGAGACCAATGTGCTTGGAGGACAGCATGTGGAGA 2822  
Db 2760 TGCACCAAGTGTCTAAGAGTGGCGGAGACCAATGTGCTTGGAGGACAGCATGTGGAGA 2819  
QY 2823 TTGCACGAGCGCAGCTCTGAGTGCATGTGTGAGCAACATGAAGCAGTGTGTGGACTC 2882  
Db 2820 TTGCACGAGCGCAGCTCTGAGTGCATGTGTGAGCAACATGAAGCAGTGTGTGGACTC 2879  
QY 2883 CAATGCTATGTGGCTCTCTTCCCTTTGGCCAGTGTGAATGGTATACGATGAGCAC 2942  
Db 2880 CAATGCTATGTGGCTCTCTTCCCTTTGGCCAGTGTGAATGGTATACGATGAGCAC 2939  
QY 2943 CTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCTATCTTGGAGCAACAGG 3002  
Db 2940 CTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCTATCTTGGAGCAACAGG 2999  
QY 3003 CTGTGCTGTGTACTGTATCCAGCAATCTGCGCAAGGGAATTCATAGAGGTTCTTA 3062  
Db 3000 CTGTGCTGTGTACTGTATCCAGCAATCTGCGCAAGGGAATTCATAGAGGTTCTTA 3059  
QY 3063 TAAAGGACAGTGAAGTGTCTTCCAGGCGCTTACAGCAATTTCTATCCACAGGCCCT 3122  
Db 3060 TAAAGGACAGTGAAGTGTCTTCCAGGCGCTTACAGCAATTTCTATCCACAGGCCCT 3119  
QY 3123 GCTCAATTCACAGTGTCTTAGAGGACAGAGATACAACTGGTCTTCAATTCAGTCTCC 3182  
Db 3120 GCTCAATTCACAGTGTCTTAGAGGACAGAGATACAACTGGTCTTCAATTCAGTCTCC 3179  
QY 3183 AGCTTGCCCAATCAACGGCCACAGTAATATGATCAATCAGAGCATCTGTGAGAGTGTGA 3242  
Db 3180 AGCTTGCCCAATCAACGGCCACAGTAATATGATCAATCAGAGCATCTGTGAGAGTGTGA 3239  
QY 3243 GAACTGACCAAGGACAGTGTGAGGACCTGATATCTGCTTCTAGGTCATCCAC 3302  
Db 3240 GAACTGACCAAGGACAGTGTGAGGACCTGATATCTGCTTCTAGGTCATCCAC 3299  
QY 3303 CAATGAGGGAATGTGAGGCAATGCAAGTGGCAAGCGCTCTCTGTGCAACACCAA 3362  
Db 3300 CAATGAGGGAATGTGAGGCAATGCAAGTGGCAAGCGCTCTCTGTGCAACACCAA 3359  
QY 3363 CAGGGCAAGTGTCTGCAACCAAGGCGCTCAAGGGGACAGTGTGATGTGA 3422  
Db 3360 CAGGGCAAGTGTCTGCAACCAAGGCGCTCAAGGGGACAGTGTGATGTGA 3419  
QY 3423 GGTAGAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTTATATCTCTTCTTAT 3482  
Db 3420 GGTAGAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTTATATCTCTTCTTAT 3479  
QY 3483 TGAATATCAGTTCACCTTTAGTCTATCCAGGAAGATGATGCTATTATACAGACTATCAA 3542  
Db 3480 TGAATATCAGTTCACCTTTAGTCTATCCAGGAAGATGATGCTATTATACAGACTATCAA 3539  
QY 3543 TTTTGTGGTACTCTCTGACGAAACAAACAGGATTTGGACATGTTTCAATAGCTCTCAA 3602  
Db 3540 TTTTGTGGTACTCTCTGACGAAACAAACAGGATTTGGACATGTTTCAATAGCTCTCAA 3599  
QY 3603 GAATTTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGA 3662

Db 3600 GAATTTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGA 3659  
QY 3663 AGAGATGCTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGA 3722  
Db 3660 AGAGATGCTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGA 3719  
QY 3723 GAAGTTTGAATTTTCCAAACCAACCAATATCACTTCTTTGTTATGTAGTAAATTTTAC 3782  
Db 3720 GAAGTTTGAATTTTCCAAACCAACCAATATCACTTCTTTGTTATGTAGTAAATTTTAC 3779  
QY 3783 CTGGCCCATCAAAATTCAGGTCAAACTGAACAAATGA 3819  
Db 3780 CTGGCCCATCAAAATTCAGGTCAAACTGAACAAATGA 3816

RESULT 2  
US-09-245-041-14  
; Sequence 14, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; EARLIER FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 8589  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-245-041-14

Query Match 96.8%; Score 3696.2; DB 3; Length 8589;  
Best Local Similarity 98.5%; Pred. No. 0;  
Matches 3744; Conservative 0; Mismatches 48; Indels 9; Gaps 1;

QY 3 GGTGGCGCAGCGCGCAACTGAGGCAAGCTGAGGAGGAGGACCGCGCGCAGCGCAGC 62  
Db 9 GGTGGCGCAGCGCGCAACTGAGGAGGAGGACCGCGCGCAGCGCAGC 68  
QY 63 GCTCGCGGCGAGGAGC 122  
Db 69 GCTCGCGGCGAGGAGC 128  
QY 123 GGGGCTGGGGC 182  
Db 129 GGGACCGC 188  
QY 183 GCTCTGCTGCTGTGTTGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 242  
Db 189 GCTCTGCTGCTGTCTT-----TTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 239  
QY 243 CGAGGC 302  
Db 240 CGAGGCGCGCTGCGGTGGC 299  
QY 303 TGACCGCGCGCTGTCAACGCGCGGTGCTGCAACCCCTGGCACCCCGCGCGCGCGCGCGC 362  
Db 300 TGAACCGCGCGGTGTCAACGCGCGGTGCTGCAACCCCTGGCACCCCGCGCGCGCGCGC 359  
QY 363 CGCGCGCTGGGTGGCGAGCAATGCCAGCACTGCGGGGCGCGCTTACAGTAACTGATC 422  
Db 360 CGCGCGCTGGGTGGCGAGCAATGCCAGCACTGCGGGGCGCGCTTACAGTAACTGATC 419  
QY 423 TCTGGGTTTGTGACAGATGGACCTGGAAATTAATAATACAAACGAAGTGCAGCTGGCT 482

Db 420 TTCTGGGTTTGTGACAGATGGACCTCGAAATTAATAACAAACGAAGTGCACGTGGCT 479  
QY 483 CATTGAAGGACAGCCAAATAGATAATGAGACTTCGTTTCATCATTTTGTCTACAGAGTG 542  
Db 480 CATTGAAGGACAGCCAAATAGATAATGAGACTTCGTTTCATCATTTTGTCTACAGAGTG 539  
QY 543 TAGTTGGGACCAATTTATATGTTTATGATGGGGACTCAATTTATGACCCCTAGTTGCTGC 602  
Db 540 TAGTTGGGACCAATTTATATGTTTATGATGGGGACTCAATTTATGACCCCTAGTTGCTGC 599  
QY 603 ATTATAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTC 662  
Db 600 ATTATAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTC 659  
QY 663 CACATCAGGTTATGCTTGTGCTGCAATTTTATGATGCTGCTTATTAATTTACATGGATT 722  
Db 660 CACATCAGGTTATGCTTGTGCTGCAATTTTATGATGCTGCTTATTAATTTACATGGATT 719  
QY 723 TAATATTAATTAAGTTTATGATGCTGCTGCAATTTTATGATGCTGCTTATTAATTTACATGGATT 782  
Db 720 TAATATTAATTAAGTTTATGATGCTGCTGCAATTTTATGATGCTGCTTATTAATTTACATGGATT 779  
QY 783 GATCAGTAATAGCAGCGAAACTGTTGAAATGTAATGTTCTGAAACTGGAAGGTGAAGC 842  
Db 780 GATCAGTAATAGCAGCGAAACTGTTGAAATGTAATGTTCTGAAACTGGAAGGTGAAGC 839  
QY 843 ATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCTCATCGAGGCACTCTGCAATTC 902  
Db 840 ATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCTCATCGAGGCACTCTGCAATTC 899  
QY 903 AAGTGAATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAGGCTCTGATGTTTCAGTTCC 962  
Db 900 AAGTGAATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAGGCTCTGATGTTTCAGTTCC 959  
QY 963 TGTACCACTAACCACTGTTTGGACTCGAGGAAATTTCTAACTTAAAGCTCCCCAG 1022  
Db 960 TGTACCACTAACCACTGTTTGGACTCGAGGAAATTTCTAACTTAAAGCTCCCCAG 1019  
QY 1023 AGCATCTTAAGCTGTGGTCAATGGAAACATTAATGTTGGTTTGTGGAGGATATATTT 1082  
Db 1020 AGCATCTTAAGCTGTGGTCAATGGAAACATTAATGTTGGTTTGTGGAGGATATATTT 1079  
QY 1083 CAACCACTCAGATTAATACATGTTTCTAGCGPATCAGCTTCTCTAGGAGTGGCTTCC 1142  
Db 1080 CAACCACTCAGATTAATACATGTTTCTAGCGPATCAGCTTCTCTAGGAGTGGCTTCC 1139  
QY 1143 ACTAAACCGTTCTGTGAACAATGTGGTTGTAGATATGTTGTCATTTCTTTGGCAATTATCAA 1202  
Db 1140 ACTAAACCGTTCTGTGAACAATGTGGTTGTAGATATGTTGTCATTTCTTTGGCAATTATCAA 1199  
QY 1203 GATAAATTTACATGTATGGAGGAAATTTGATCCAATCGGATGTGACCAATGAGTT 1262  
Db 1200 GATAAATTTACATGTATGGAGGAAATTTGATCCAATCGGATGTGACCAATGAGTT 1259  
QY 1263 GAGAGTTTTTCAATTCATATGATGCTATGGTGTGTTGACCCCTTAAGGCAAAAGGAGCA 1322  
Db 1260 GAGAGTTTTTCAATTCATATGATGCTATGGTGTGTTGACCCCTTAAGGCAAAAGGAGCA 1319  
QY 1323 GTATCAGTGGTTGGGCACTGTGCAACATTTGTTACATGAAGAAATGGCCGAGTGGTCAT 1382  
Db 1320 GTATCAGTGGTTGGGCACTGTGCAACATTTGTTACATGAAGAAATGGCCGAGTGGTCAT 1379  
QY 1383 GCTGTGCTATCTTTGGTCTACTGCTCTCTATGGATATATAGCAATGTGAGGAAATATGA 1442  
Db 1380 GCTGTGCTATCTTTGGTCTACTGCTCTCTATGGATATATAGCAATGTGAGGAAATATGA 1439  
QY 1443 TTTGATTAAGAACACATCGAGTATATTAACACCCAGGGTCCCTTGTGCAAGGGGGTTA 1502  
Db 1440 TTTGATTAAGAACACATCGAGTATATTAACACCCAGGGTCCCTTGTGCAAGGGGGTTA 1499  
QY 1503 CGGCCATAGCAGTGTTTTACGACCATAGGACAGGCGCTTATACGTTTCATGTGCTACAA 1562

Db 1500 CGGCCATAGCAGTGTTTTACGACCATAGGACAGGCGCCCTATACGTTTCATGTGGTACAA 1559  
QY 1563 GGCTTTTCACTGCGCAATTAAGTATACCGGCTTGAGATGATCTCTACCGATATGATGTGGATAC 1622  
Db 1560 GGCTTTTCACTGCGCAATTAAGTATACCGGCTTGAGATGATCTCTACCGATATGATGTGGATAC 1619  
QY 1623 CCAGATGTGGACCACTTCTTAAGGACAGCCGATTTTCCGTTTACTTGCACACAGCTGTGAT 1682  
Db 1620 CCAGATGTGGACCACTTCTTAAGGACAGCCGATTTTCCGTTTACTTGCACACAGCTGTGAT 1679  
QY 1683 AGTGAGTGGAAACCACTGCTGTGTTGGGGGAAACACACAAATGACACATCTATGAGCCA 1742  
Db 1680 AGTGAGTGGAAACCACTGCTGTGTTGGGGGAAACACACAAATGACACATCTATGAGCCA 1739  
QY 1743 TGGCGCCAAATGCTTCTTTCAGATTTTCATGGGCTATGACATTCCTCTCAGCAGCTGCTC 1802  
Db 1740 TGGCGCCAAATGCTTCTTTCAGATTTTCATGGGCTATGACATTCCTCTCAGCAGCTGCTC 1799  
QY 1803 AGTGCTTCCAGACCTGATCTCCACCAATGATGTCACAGATTTTGGCCATTTCAGCAGTCTT 1862  
Db 1800 AGTGCTTCCAGACCTGATCTCCACCAATGATGTCACAGATTTTGGCCATTTCAGCAGTCTT 1859  
QY 1863 ACACAAACAGCAGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1922  
Db 1860 ACACAAACAGCAGCACTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1919  
QY 1923 GGTATTCACCTCGGACAGTGTGATCGCATCGAGTGAAGCGCTTGTTTAGCAGCAGG 1982  
Db 1920 GGTATTCACCTCGGACAGTGTGATCGCATCGAGTGAAGCGCTTGTTTAGCAGCAGG 1979  
QY 1983 ACCTGTTATTCGGTGTGTGTGGAACACAGGTCGTCTCAGTGTATCTCTGTTGGCGCTGGC 2042  
Db 1980 ACCTGTTATTCGGTGTGTGTGGAACACAGGTCGTCTCAGTGTATCTCTGTTGGCGCTGGC 2039  
QY 2043 AACTGTATGAACAAGAAAGTTTAAATCAGAAATGTTTTTCCAAAGAACTCTTGACCA 2102  
Db 2040 AACTGTATGAACAAGAAAGTTTAAATCAGAAATGTTTTTCCAAAGAACTCTTGACCA 2099  
QY 2103 TGACAGATGTGACCAACACACAGATTTTACAGCTCTACAGCCAAACACCAATGACTGCCA 2162  
Db 2100 TGACAGATGTGACCAACACACAGATTTTACAGCTCTACAGCCAAACACCAATGACTGCCA 2159  
QY 2163 CTGGTCAATGACCAATGTTGTTCCCGAGGAAACACAGCTGCTCAGAGGCGCAGATCTCCAT 2222  
Db 2160 CTGGTCAATGACCAATGTTGTTCCCGAGGAAACACAGCTGCTCAGAGGCGCAGATCTCCAT 2219  
QY 2223 TTTTAGGTATGAGATTTGCCCAAGGATAACCCCATGTACTACTGTAAACAAAGAACCCAG 2282  
Db 2220 TTTTAGGTATGAGATTTGCCCAAGGATAACCCCATGTACTACTGTAAACAAAGAACCCAG 2279  
QY 2283 CTGACAGGCTGTGCTCTGGAACAGAACTGCGAGTGGGAGCCCGGAAATCAGAGTGTGAT 2342  
Db 2280 CTGACAGGCTGTGCTCTGGAACAGAACTGCGAGTGGGAGCCCGGAAATCAGAGTGTGAT 2339  
QY 2343 TGCCCTTCCCGGAAAATATCTGTGGCAATTCGCTGGCAATTTGGTTGAAACTCAITGTTGAA 2402  
Db 2340 TGCCCTTCCCGGAAAATATCTGTGGCAATTCGCTGGCAATTTGGTTGAAACTCAITGTTGAA 2399  
QY 2403 AATTACTCTGCGAAGGAAATTTAGCAATGCTAAATTTGTTCTGTAGGAAACCAATGC 2462  
Db 2400 AATTACTCTGCGAAGGAAATTTAGCAATGCTAAATTTGTTCTGTAGGAAACCAATGC 2459  
QY 2463 CTTTGTGGCTTCTCTTACAAACCCAGAGAGGTAGAAATTTGCTTTAAGCAGCTCGGAT 2522  
Db 2460 CTTTGTGGCTTCTCTTACAAACCCAGAGAGGTAGAAATTTGCTTTAAGCAGCTCGGAT 2519  
QY 2523 AATGAGTCTATCTCAGAGCATGTCAAGCTCACTTAAACCCCATGGTTCGGCTTCGGAA 2582  
Db 2520 AATGAGTCTATCTCAGAGCATGTCAAGCTCACTTAAACCCCATGGTTCGGCTTCGGAA 2579  
QY 2583 GATCATGTGCTCTACTGCTGCGAGAGATATGTCCTTCAATTTACAAATAGTTTACTACA 2642  
Db 2580 GATCATGTGCTCTACTGCTGCGAGAGATATGTCCTTCAATTTACAAATAGTTTACTACA 2639

QY 2643 GTGATGCGCTCTGAGCCAGTGTATCTGGAATCTGTGAAATTTTATCAGAACCCAGTAC 2702  
Db 2640 GTGATGCGCTCTGAGCCAGTGTATCTGGAATCTGTGGAATTTTATCAGAACCCAGTAC 2699  
QY 2703 TCGGGAGCTGAAGGCTGCAACCTGTGATCAACCCACTCAATGTGTAGTCTGTGAAAGGCC 2762  
Db 2700 TCGGGAGCTGAAGGCTGCAACCTGTGATCAACCCACTCAATGTGTAGTCTGTGAAAGGCC 2759  
QY 2763 TGCARACACAGTCTAAGGAGTCCGGACACATGTGCTTGAGACAGCATGTGGAGA 2822  
Db 2760 TGCARACACAGTCTAAGGAGTCCGGACACATGTGCTTGAGACAGCATGTGGAGA 2819  
QY 2823 TTGACACAGCGGAGCTCTGAGTGCATGTGTGCGACACATCAAGACAGTGTGTGCACTC 2882  
Db 2820 TTGACACAGCGGAGCTCTGAGTGCATGTGTGCGACACATCAAGACAGTGTGTGCACTC 2879  
QY 2883 CAATGCTATGTGGCTCTCCCTTTTGGCCAGTGTATGGATGGTATACATGAGCAC 2942  
Db 2880 CAATGCTATGTGGCTCTCCCTTTTGGCCAGTGTATGGATGGTATACATGAGCAC 2939  
QY 2943 CTGCCCCCTGAAATTTGTTAGGCTACTGTACTCTGTAGTCTATGTTGGAGCAACCAAG 3002  
Db 2940 CTGCCCCCTGAAATTTGTTAGGCTACTGTACTCTGTAGTCTATGTTGGAGCAACCAAG 2999  
QY 3003 CTGTGCTGTGTACTGATCCCGAATACTGCGAAGGAAATGCATAGAGGTTCTCTA 3062  
Db 3000 CTGTGCTGTGTACTGATCCCGAATACTGCGAAGGAAATGCATAGAGGTTCTCTA 3059  
QY 3063 TAAAGGACAGTCAAGATGCTTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCT 3122  
Db 3060 TAAAGGACAGTCAAGATGCTTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCT 3119  
QY 3123 GCTCAATTCAGCATGTGTAGAGACAGAGATACAACTGGTCTTTCTATCTACTGTCC 3182  
Db 3120 GCTCAATTCAGCATGTGTAGAGACAGAGATACAACTGGTCTTTCTATCTACTGTCC 3179  
QY 3183 AGCTTGCAATGCAAGGACACAGTAATGCATCAATCAGACATCTGTGAGAGTGTGA 3242  
Db 3180 AGCTTGCAATGCAAGGACACAGTAATGCATCAATCAGACATCTGTGAGAGTGTGA 3239  
QY 3243 GAACCTGACACAGGCAAGCACTGCGAGACCTGCAATCTGCTTCTACGGTGTATCCAC 3302  
Db 3240 GAACCTGACACAGGCAAGCACTGCGAGACCTGCAATCTGCTTCTACGGTGTATCCAC 3299  
QY 3303 CAATGAGGGAATGTGACGATGCAAGTGCATGAGGACGCTCTGTGTGCAACACCAA 3362  
Db 3300 CAATGAGGGAATGTGACGATGCAAGTGCATGAGGACGCTCTGTGTGCAACACCAA 3359  
QY 3363 CACGGCAAGTGTCTTGCAACCAAGGCGGTCAAGGGGACGAGTGCAGCTATGTGA 3422  
Db 3360 CACGGCAAGTGTCTTGCAACCAAGGCGGTCAAGGGGACGAGTGCAGCTATGTGA 3419  
QY 3423 GGTAGAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTTATATCTTCTTTAT 3482  
Db 3420 GGTAGAAATTCGATACCAAGGAACCTCTCAGAGGAACATGTTATATCTTCTTTAT 3479  
QY 3483 TGAATATCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATCAA 3542  
Db 3480 TGAATATCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATCAA 3539  
QY 3543 TTTTGTGGCTACTCTGACGACAAACAGGATTTGGACATGTTTCATCATGCTCCAA 3602  
Db 3540 TTTTGTGGCTACTCTGACGACAAACAGGATTTGGACATGTTTCATCATGCTCCAA 3599  
QY 3603 GAATTTCAACCTCAACATCACTTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3662  
Db 3600 GAATTTCAACCTCAACATCACTTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGA 3659  
QY 3663 AGAGATGCTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATCA 3722  
Db 3660 AGAGATGCTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATCA 3719

QY 3723 GAAGTTTGATTTTCGCAACCAACCCAAATATCACTTCTTTGTTTATGTCAAGTAATTCAC 3782  
Db 3720 GAAGTTTGATTTTCGCAACCAACCCAAATATCACTTCTTTGTTTATGTCAAGTAATTCAC 3779  
QY 3783 CTGCCCCATCAAAATTCAGGT 3803  
Db 3780 CTGCCCCATCAAAATTCAGAT 3800

RESULT 3  
US-09-245-041-1  
; Sequence 1, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; EARLIER FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 8827  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-1

Query Match 77.0%; Score 2939.8; DB 3; Length 8827;  
Best Local Similarity 85.1%; Pred. No. 0;  
Matches 3400; Conservative 0; Mismatches 392; Indels 205; Gaps 2;  
QY 3 GGTGGCGGCGAGCGGCGCAACTGAGCAAGGCTGAGGAGGAGGACGGGGGACCGGAGC 62  
Db 82 GGTGGCGGCGGCGGCGCACTGAGGCGGCTGAGGGGAGGACGAGGACGAGCAGC 141  
QY 63 GCTCGGGGAGAGCGGCGGCGGCACTGAGGAGCTGAGGAGCTGAGGAGGCTGGAGGCC 122  
Db 142 GCCTGGGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 201  
QY 123 GGGGCTGGGGGCGGCGGCTCGGCTCCCGGGCTGTGTCTCCACCGCTGCGGCCACGGCT 182  
Db 202 GGGACCGGCGGCGGCGGCTGTGTCTCCCGGGGCTGTGTCTCGGGGCGCTGCCCGGCC 261  
QY 183 GCTGCTGCTGCTGTGTGTCTCCCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 242  
Db 262 GCTGCTGCTGCTGTCTT-----TTGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 312  
QY 243 CGAGGCGGCGGCGGCGGCGGCGGCTGTGCGGCTCAGCGCAGCGGAGGCGGCGGAGGATG 302  
Db 313 CGAGGCGGCTGT 372  
QY 303 TGACCGGCGGCTGT 362  
Db 373 TGACCGGCGGCTGT 432  
QY 363 CGCGGCTGGTGGGCGGAGCAATGCGACATGCGGGGCGGCTTCAG-----410  
Db 433 CACGGCTGGTGGGCGGAGCAATGCGACATGCGGGGCGGCTTCAGGACATCTGTGTCTC 492  
QY 411 -----410  
Db 493 ACGCTTAATACAGCTGTTCGGAAGGTGAGGCTGGAGGAACAGTTCGAGGAAGCTTC 552  
QY 411 -----410  
Db 553 GGCTACAGAAATAGTTCAAGAGTAACCTGGGGCAACTTGGGCTTGTCTTCCAAACCAAAA 612

411 ----- 410  
Db 613 TGACGAAAGGACAGCTAGAGTCTTTTGGGAAATTTTACGTGACTAAATTTTCACC 672  
Qy 411 ----- ACTAATCTGATCTCTGGGTTGTGACAGATGGACCTGGAAATTTATAATACAAA 466  
Db 673 GAGAACTAACTGGCTCTCTCGAATTTGTAACAGATGGACCTGGGAATTTATAATATAAGA 732  
Qy 467 CGAAGTGCAGCTGCTCATTTAGAGGACAGCCAAATAGAAATAGACATCTGTTCAATC 526  
Db 733 CGAAGTGCAGCTGCTCATTTAGAGGACAGCCAAATAGAAATAGACATCTGTTCAATC 792  
Qy 527 ATTTTGTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATG 586  
Db 793 ATTTTGTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATG 852  
Qy 587 CACCGTAGTTGCTGATTTAGTTGGCTCATTTGTTCTGCTGCAATTTTATGATGGGACTG 646  
Db 853 CACCTCTGATTTGCTGCTTTAGTGGCTCATTTGTTCTGAAAGAGATGGCAATGAGACGG 912  
Qy 647 TCCTCAGAGTTGTTGCCACATCAGGTTATGCTTGTCTGCAATTTTATGATGGGACTCAATTTATG 706  
Db 913 CTCTCAGAGTCACTGTCACTTCAGTTATGCACTGCTGCAATTTTATGATGGGACTGCTT 972  
Qy 707 ATAAATTTGACTGGAATTAATATACCTACAGTTTGTATGTTGATGTTGTCCTGCTCAG 766  
Db 973 ATAAATCTGACTGGAATTAATATCACTTACAAATTTTGACATGTTCCGAATTAATCTCAG 1032  
Qy 767 GCGGAGGAGTGTAGATCAGTATAGCAGGAACTGTTCAATGTAATGTTCTGAAA 826  
Db 1033 GCGGAGGAGTGTAGACAGTATACAGCAGGCTGTTGAGTGTGAATGTTCTGAAA 1092  
Qy 827 ACTGGAAGGTTGAAGCATGTGACATCTCTCACTGTACAGACAACTGTTGCTCATC 886  
Db 1093 ACTGGAAGGAGTGTGCTGTGACATCTCTCACTGTACAGACAACTGTTGCTCATC 1152  
Qy 887 GAGGATCTGCAATTTCAAGTGTATGTCAGAGATGCTCTGTTCTCAGACTGGCAGGCT 946  
Db 1153 GAGGATCTGTAATGCAAGCGATACAGAGGCTGCTCTGCTTCTCACTGGCAGGCT 1212  
Qy 947 CTGGATGTTCACTCTCTGACAGCTTAACAGTCAATTTTGACTGAGAGAAATTTCTA 1006  
Db 1213 CTGGATGTTCAATTTCTGCTGACCTTACCAGTCTTTTGGACTGAGAGAAATTTCTG 1272  
Qy 1007 ACTTAAAGTCTCCAGAGCACTCTAAAGCTGTGGTCAATGGAACAAATTTATGCGGTTG 1056  
Db 1273 ATTTAAAGCTCTCCAGAGCTCTCAATAAGCTGTGGTCAATGGAACAAATTTATGCGGTTG 1332  
Qy 1067 TTGGAGGATATATGTTCAACCACTCAGATTTATACATGTTTCTAGCGTATGACCTTCTT 1126  
Db 1333 TTGGCGGATATATGTTCAACCACTCAGATTTATACATGTTTCTAGCGTATGACCTTCTT 1392  
Qy 1127 CTAGGAGTGTCTTCCACTAAACCGTCTCTGGAACAAATGTTGTTGTAGATATGCTCAAT 1186  
Db 1393 CTAGGGAATGCTTCCACTAAACCAATCTGTGAAACAGTGTGTTGTGAAGATATGCTCAAT 1452  
Qy 1187 CTTTGGCAATTTACAGAGTAAATTTACATGTTATGAGGAAATTTGATCCAACTGGGA 1246  
Db 1453 CTTTGGCAATTTACAGAGTAAATTTACATGTTATGAGGAAATTTGATCCAACTGGGA 1512  
Qy 1247 ATGTGACCAATGAGTTGAGAGTTTTTCACATTTCAATATGAGTCAATGTTGTTGACCC 1306  
Db 1513 ACGTGACCAATGAGTTGAGAGTATTTTCATATTTCAATATGATCATGAGTATTTGTTAATC 1572  
Qy 1307 CTAAGGCAAGGAGGAGTATGCTAGTGTGGGCACTCTGACACATTTGTTTACACTGAAGA 1366  
Db 1573 CGAAGCTTAAGGATCAGTATGCTAGTGTGGGCACTCTGACACATTTGTTTACACTGGCAT 1632  
Qy 1367 ATGCGCGAGTGTGATGCTGCTCATCTTTGGTCACTGCTCCCTCTCTATGAGATATATAAGCA 1426  
Db 1633 CTGCGCGTGTGATGTTGTTGCTCATCTTCGTCATTTCCCTCACTTCCCTCACTTATGAGCG 1692  
Qy 1427 ATGTGAGGAATATGATTTGGATGAAGAACACATGGAGTATATTACACACCCAGGTTGCC 1486

Db 1693 TTGTGAGGAATATGACTTGGAAAGAAACACATGAGTATATTACATACTCAGGGTCTCT 1752  
Qy 1487 TTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTTACGACCATAGGACACAGGGCCCTATACG 1546  
Db 1753 TTGTGCAAGGGGGTTATGGCCACAGTATGTTTATGATGACAGGACCAAGGCTCTGTACG 1812  
Qy 1547 TTCTATGTTGGCTACAAGGCTTTTCACTGCCAATAGTACCGGCTTGCAGATGATCTCTACC 1606  
Db 1813 TTCTATGTTGGCTACAAGGCTTTTCACTGCCAATAGTACCGGCTTGCAGATGATCTCTACA 1872  
Qy 1607 GATATGATGTGATACCCAGATGTGGACCATCTTAAAGGACAGCCGATTTTCCGTTACT 1666  
Db 1873 GATATGATGTGATACCCAGATGTGGACCATCTTAAAGGACAGCCGATTTTCCGTTACT 1932  
Qy 1667 TGCAACAGCTGTGATAGTGTGAAACCATGCTGTGTTTGGGGGAAACACACAAATG 1726  
Db 1933 TGCAACAGCTGTGATAGTGTGAAACCATGCTGTGTTTGGGGGAAACACACAAATG 1992  
Qy 1727 ACACATCTATGAGCCATGCGCCCAATGCTTCTTCCAGATTTTCATGCGCCTATGACATG 1786  
Db 1993 ACATTTCCATGAGCCAGGTCGCAAAATGCTTCTCTCGGACTTTCATGGCTTATGACATG 2052  
Qy 1787 CCTGTGACCTGCTGCTGCTGCTTCCAGACCTTCAATGCTTCTTCCAGATTTTCATGAGATTTG 1846  
Db 2053 CTGTGACCGATGCTGCTGCTTCCAGACCTTCCAGATTTTCATGATGTCACAGATTTG 2112  
Qy 1847 GCAATTCAGAGTCTTACACACAGCAGCATGATGTTTGGTGTGTTTCAATGATCTCTC 1906  
Db 2113 GCAATTCAGAGTCTTGTACACAGCAGCAGTATGTTTGGGCGCTTTCACAGCGCTCC 2172  
Qy 1907 TCCTCAGCAGCATCTCTGCTTATTCACCTCGGAAACAGTGTGATGCGCATCGAGTGAAGCG 1966  
Db 2173 TCCTCAGTGAAGTCTTGGTCTTCTTACCTCGAGCAGTGCATGACACCCAGTGAAGCTG 2232  
Qy 1967 CTTGTTTAGCAGCAGCACTGCTGATTTTGGTGTGTTGGAACACAGGCTCCTCTCAGTGTG 2026  
Db 2233 CTTGTTGAGCAGCAGCACTGCTGATTCGGTGTCTGTGGGACACACAGTCTCTCGATGTA 2292  
Qy 2027 TCTCTGTTGGGCTGCGCACTGATGACACAGAGAAAGTAAATCAGATGTTTTCCTCA 2086  
Db 2293 CTTCTCTGGAGTGTGGCACTGAGAAACAGAGAAAGTAAATCAGAGTGTCTTCTCA 2352  
Qy 2087 AAAGAACTCTTGAACCATGACAGATGTGACAGCAGACACAGATTTTACAGCTGTACAGCA 2146  
Db 2353 AAAGAACTCTTGAACCATGACAGATGTGACAGCAGACACAGATTTTACAGCTGTACAGCA 2412  
Qy 2147 ACACCAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2206  
Db 2413 ATACCAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2472  
Qy 2207 AAGCCAGATCTCTCAATTTTATGATGATGAAATTTGCTGCTGCTGCTGCTGCTGCTGCTG 2266  
Db 2473 AAGCCAGATCTCTCAATTTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2532  
Qy 2267 GTTAAAGAGACAGCTGCTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2326  
Db 2533 GCAATTAAGAAACAGCTGCTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2592  
Qy 2327 GCAATCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2386  
Db 2593 GCAATCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2652  
Qy 2387 GAAATCTATGTTTGAATTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2446  
Db 2653 GAAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2712  
Qy 2447 GTAGGAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2506  
Db 2713 GTAGGAACCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2772  
Qy 2507 TTAAGCAGCTGCAATTAATGAGTCTATCTCAGAGCATGCTCAAGCTCACTTAAACCCCAT 2566

2773 TTAAAGCAGCTTCGATTAAATGCAATCATCTCAAAGTATGTCCAAGCTCACTCTGACTCCAT 2832  
2567 GGCTCGCCCTTCGGAAGATCAATGTGTCTCTACTGTGTGTGGGAAGATATGTCCCATTTTA 2626  
2833 GGCTTGTCTTCGGAAGATCAATGTGTCTCTACTGTGTGTGGGAAGATATGTCTCCATTTCA 2892  
2627 CAATATGTTTACTACAGTGGATGCGCTCTGAGCCAGTGTGATGCTGTGGAATTT 2686  
2893 CAATATGTTTGTGCAAGTGGATGCCATCTGAGCCAGTGTGATGCTGTGGAATTT 2952  
2687 TATCAGAACCCAGTACTCTGGGAGTGAAGCTGTCAACCTGTGATCAACCCACTCAATGTGA 2746  
2953 TGTACAGGCTAGTACTCTGGGAGTAAAGGCTGCAACCTGTGATCAACCTCTCAATGGCA 3012  
2747 GTCTCTGTGAAGGCTCTGAACCCAGTGTCTAAAGCTGTGCGGACACATGTGCTTGA 2806  
3013 GCCTCTGTGAAGGCTCTGAACCCAGTGTGCGGACACATGTGCTTGA 3072  
2807 GGACAGCATGTGAGATTTGACACAGCGGAGCTCTGAGTGTGATGTGTGCGAGCAACATGA 2866  
3073 GGACAGCATGTGAGATTTGACACAGCGGAGCTCTGAGTGTGATGTGTGCGAGTAAATGA 3132  
2867 AGCAGTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGAT 2926  
3133 AGCAGTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGAT 3192  
2927 GGTATACGATGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGAT 2986  
3193 GGTATACGATGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGAT 3252  
2987 GCTTGGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAG 3046  
3253 GCTTGGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAG 3312  
3047 GCATAGAGGCTTCTATAAGGACAGTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAG 3106  
3313 GTATTGGGCGAGTATTAAGGACCTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAG 3372  
3107 TCTATCCAGGCTTCTATAAGGACAGTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAG 3166  
3373 TGTATCCAGGCTTCTATAAGGACAGTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAG 3432  
3167 CTATTATCTCTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGAT 3226  
3433 CTATTATCTCTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGAT 3492  
3227 TCTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAG 3286  
3493 TCTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAG 3552  
3287 TCTACGCTGATCCCAATGAGGAGTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAG 3346  
3553 TCTATGCTGACCTATTAAGGACCTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAG 3612  
3347 CTCTGTGCAACCAACAGCGGAGTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGAT 3406  
3613 CACTGTGCAACCAACAGCGGAGTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGAT 3672  
3407 AGTGCCAGTGTGAGTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAG 3466  
3673 AGTGCCAGTGTGAGTGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAGTGTGATGTGAG 3732  
3467 ATTATCTCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3526  
3733 ACTATACCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3792  
3527 ATTACAGCTATCAATTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3586  
3793 ACTACAGCTATCAATTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3852  
3587 TCATCAATGCTTCAAGAAATTTCAACCTCAACATCAACCTGAGTGTGATGATGATGATGATGATGAT 3646  
3853 TCATCAATGCTTCAAGAAATTTCAACCTCAACATCAACCTGAGTGTGATGATGATGATGATGATGAT 3912

3647 GAACCCAGCTGGAGAGAGATGCTGTGTTTCAAAACCAACATTAAGAGTACAAAG 3706  
3913 GAACCCAGCTGGAGAGAGATGCTGTGTTTCAAAACCAACATTAAGAGTACAAAG 3972  
3707 ATAGTCTCTCTAATGAGATTTGATTTTTCGCAACCCCAACATTAACATTTCTTTT 3766  
3973 ATAGTCTCTCTAATGAGATTTGATTTTTCGCAACCCCAACATTAACATTTCTTTT 4032  
3767 ATGTCAAGTAAATTCACCTGGCCCATCAAAATTCAGGT 3803  
4033 ATGTCAAGTAAATTCACCTGGCCCATCAAAATTCAGAT 4069

RESULT 4  
US-09-245-041-18  
; Sequence 18, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 2625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-245-041-18

Query Match 58.9%; Score 2248.2; DB 3; Length 2625;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 2308; Conservative 0; Mismatches 68; Indels 9; Gaps 1;

3 GGTGGCCGCGAGCGCGGCAACTGAGGCAAGGCTGAGGAGGAGGAGCGGCGGCGAGCGGAGC 62  
9 GGTGGCCGCGAGCGCGGCAACTGAGGCAAGGCTGAGGAGGAGGAGCGGCGGCGAGCGGAGC 68  
63 GCTCGCGGCGAGGAGCGGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGAGCGGCGGCGAGCGG 122  
69 GCTCGCGGCGAGGAGCGGCGGCGGCAACTGAGGCAAGGCTGAGGAGGAGGAGCGGCGGCGAGCGG 128  
123 GGGGCTTGGGGCGCGGCGGCTGCGCTCCCGCGGCTGCTCTCCCGCGGCGGCTGCGCGGCGGCGG 182  
129 GGGACCGCGCGCGCGGCTGCTCTCCCGCGGCGGCTGCTCTCCCGCGGCGGCTGCGCGGCGGCGG 188  
183 GCTG 242  
189 GCT 239  
243 CGAGGCGCGGCGGCGGCGGCGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCGGCGGCGGCGG 302  
240 CGAGGCGCGGCGGCGGCGGCGGCGGCTGCTGCGGCTGCTGCGGCTGCTGCGGCGGCGGCGGCGG 299  
303 TGACGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 362  
300 TGACGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 359  
363 CGCGGCT 422  
360 CGCGGCT 419  
423 TTCTGCT 482  
420 TTCTGCT 479

Qy 483 CATTGAAGGACAGCCAAATAGATAATATGAGACTTCGTTTCAATCAATTTTGTCTACAGAGTG 542  
Db 480 CATTGAAGGACAGCCAAATAGATAATAGAGCTTCGTTTCAATCAATTTTGTCTACAGAGTG 539  
Qy 543 TAGTGGGACCAATATATGTTTATGATGGGACCTCAATTTATGACACCGCTAGTGTGTC 602  
Db 540 TAGTGGGACCAATATATGTTTATGATGGGACCTCAATTTATGACACCGCTAGTGTGTC 599  
Qy 603 ATTTAGTGGCCTCAATTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGC 662  
Db 600 ATTTAGTGGCCTCAATTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGC 659  
Qy 663 CACATCAGGTTATGCTGCTGCAATTTTATGATGATGCTGCTTATATATTTGACTGGATT 722  
Db 660 CACATCAGGTTATGCTGCTGCAATTTTATGATGATGCTGCTTATATATTTGACTGGATT 719  
Qy 723 TAATATTACTTACAGTTTGTATGTTGTCCTCAATTAATGCTCAGGCCGAGGAGAGTGTA 782  
Db 720 TAATATTACTTACAGTTTGTATGTTGTCCTCAATTAATGCTCAGGCCGAGGAGAGTGTA 779  
Qy 783 GATCAGTATAGCAGCAACTGTTCAATGTAATGTTCTGAAACTGGAAGTGAAGC 842  
Db 780 GATCAGTATAGCAGGAGTACTGTTGAATGTAATGTTCTGAAACTGGAAGTGAAGC 839  
Qy 843 ATGTGACATTCCTCACTGTACAGCAAACTGTGGTTTTCTCATCGAGGCATCTGCAATTC 902  
Db 840 ATGTGACATTCCTCACTGTACAGCAAACTGTGGTTTTCTCATCGAGGCATCTGCAATTC 899  
Qy 903 AAGTGTATGACAGAGTCTCCTGCTTCTCAGACTGGCAGGGTCTGATGTTCAAGTTC 962  
Db 900 AAGTGTATGACAGAGTCTCCTGCTTCTCAGACTGGCAGGGTCTGATGTTCAAGTTC 959  
Qy 963 TGTACCAGCTAACAGCTCAATTTTGGACTCGAGAGGAATATCTAACTTTAAAGCTCCCCAG 1022  
Db 960 TGTACCAGCTAACAGCTCAATTTTGGACTCGAGAGGAATATCTAACTTTAAAGCTCCCCAG 1019  
Qy 1023 AGCATCTCAATGCTGTGCTGATGGAACATTAATGTTGGTTGTTGAGGATATATGTT 1082  
Db 1020 AGCATCTCAATGCTGTGCTGATGGAACATTAATGTTGGTTGTTGAGGATATATGTT 1079  
Qy 1083 CAACCACTCAGATATATCAATGCTTACGCTATGACCTTGTCTTAGGGAGTGGCTTCC 1142  
Db 1080 CAACCACTCAGATTAATCAATGCTTACGCTATGACCTTGTCTTAGGGAGTGGCTTCC 1139  
Qy 1143 ACTAAACCGTTCTGTGAACATGTTGTTGTTAGATATGTTCACTTTTGGCATATATCAA 1202  
Db 1140 ACTAAACCGTTCTGTGAACATGTTGTTGTTAGATATGTTCACTTTTGGCATATATCAA 1199  
Qy 1203 GGTATAAAATTTACATGATGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAGTT 1262  
Db 1200 GGTATAAAATTTACATGATGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAGTT 1259  
Qy 1263 GAGAGTTTTCATATTAATGATGATGCTGTTGTTGACCTTAAGCCCAAGGAGCA 1322  
Db 1260 GAGAGTTTTCATATTAATGATGATGCTGTTGTTGACCTTAAGCCCAAGGAGCA 1319  
Qy 1323 GTATGAGTGGTGGGCACTCTGCACACATTTGTATACACTGAAGATGGCCGAGTGGTCAT 1382  
Db 1320 GTATGAGTGGTGGGCACTCTGCACACATTTGTATACACTGAAGATGGCCGAGTGGTCAT 1379  
Qy 1383 GCTGGTCATCTTTGGTCACTGCTCTCTATGGATATATAGCAATGTGCAAGATATGA 1442  
Db 1380 GCTGGTCATCTTTGGTCACTGCTCTCTATGGATATATAGCAATGTGCAAGATATGA 1439  
Qy 1443 TTTGGATAGAACAACATGAGTATATATACACCCAGGGTGGCTTGTGCAAGGGGTTA 1502  
Db 1440 TTTGGATAGAACAACATGAGTATATATACACCCAGGGTGGCTTGTGCAAGGGGTTA 1499  
Qy 1503 CGGCCATAGCAGTGTATGACCATAGGACAGGGCCCTTACGTTCAATGTGCTACAA 1562  
Db 1500 CGGCCATAGCAGTGTATGACCATAGGACAGGGCCCTTACGTTCAATGTGCTACAA 1559

Qy 1563 GGCTTTCACTGTCATTAAGTACCGCTTGCAGATGATCTCTACCGATATGATGGATAC 1622  
Db 1560 GGCTTTCACTGTCATTAAGTACCGCTTGCAGATGATCTCTACCGATATGATGGATAC 1619  
Qy 1623 CAGATGTGACCAATCTTTAAGGACAGCCGATTTTTCGTTACTTGCACACAGCTGTGAT 1682  
Db 1620 CAGATGTGACCAATCTTTAAGGACAGCCGATTTTTCGTTACTTGCACACAGCTGTGAT 1679  
Qy 1683 AGTGTGGAACCAATGCTGGTGTGTTGGGGGAAACACACAAATGACATCTATGAGCCA 1742  
Db 1680 AGTGTGGAACCAATGCTGGTGTGTTGGGGGAAACACACAAATGACATCTATGAGCCA 1739  
Qy 1743 TGGCGCAAAATGCTTCTTCAGATTTTCATGGCTTATGACATTTGCTGTGACCCCTGGTC 1802  
Db 1740 TGGCGCAAAATGCTTCTTCAGATTTTCATGGCTTATGACATTTGCTGTGACCCCTGGTC 1799  
Qy 1803 AGTGTTCCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGAGTCTT 1862  
Db 1800 AGTGTTCCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGAGTCTT 1859  
Qy 1863 ACACAAACAGCAGCATGATGTTTCGGTGTGTTCAATAGTCTCTCTCCTCAGCAGCATCCT 1922  
Db 1860 ACACAAACAGCAGCATGATGTTTCGGTGTGTTCAATAGTCTCTCTCCTCAGCAGCATCCT 1919  
Qy 1923 GGTATTACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTGTAGCAGG 1982  
Db 1920 GGTATTACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTGTAGCAGG 1979  
Qy 1983 ACTGTTATCGTGTGTTGTTGGAACACAGGGTCTGCTCAGTGTATCTCGTGGGCGCTGGC 2042  
Db 1980 ACTGTTATCGTGTGTTGTTGGAACACAGGGTCTGCTCAGTGTATCTCGTGGGCGCTGGC 2039  
Qy 2043 AACTGATGAACAGGAAAGAAAGTTAAATCAGAATGTTTTCCTCAAAAGAACTCTTGACCA 2102  
Db 2040 AACTGATGAACAGGAAAGAAAGTTAAATCAGAATGTTTTCCTCAAAAGAACTCTTGACCA 2099  
Qy 2103 TCACAGATGTGACAGCAGCAGATGTTTACAGCTGTACAGCCAAACCAATGACTGCCA 2162  
Db 2100 TCACAGATGTGACAGCAGCAGATGTTTACAGCTGTACAGCCAAACCAATGACTGCCA 2159  
Qy 2163 CTGTGCAATGACCAATGTTGTTCCCGAGGAACCAAGCTGTCTCAGAGGCCAGATCTCCAT 2222  
Db 2160 CTGTGCAATGACCAATGTTGTTCCCGAGGAACCAAGCTGTCTCAGAGGCCAGATCTCCAT 2219  
Qy 2223 TTTTGTGATGAGATTCGCCCAAGGATACCCCATGTACTGTACTGTAAAGAGACAG 2282  
Db 2220 TTTTGTGATGAGATTCGCCCAAGGATACCCCATGTACTGTACTGTAAAGAGACAG 2279  
Qy 2283 CTGAGGAGCTGTGCTCGACCAAGCTGCCAGTGGAGCCCGGAATCAGGAGTGCAT 2342  
Db 2280 CTGAGGAGCTGTGCTCGACCAAGCTGCCAGTGGAGCCCGGAATCAGGAGTGCAT 2339  
Qy 2343 TGCCCTGCCGGAATATCTGTGCAATTTGGCTGCAATTTGGTTGG 2387  
Db 2340 TGCCCTGCCGGAATATCTGTGCAATTTGGCTGCAATTTGGTTGG 2384

## RESULT 5

US-09-245-041-8  
; Sequence 8, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; OF OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20

[illegible]



Db 2053 CTTGTGACCGATGGTCTAGTCTCCAGACCTGAGCTCCATCATGATGTCACAGATTG 2112  
QY 1847 GCCATTACAGAGTCTTACACAGACGACCATGATGATGTTGGTGGTTCATATAGTCTCC 1906  
Db 2113 GCCATTACAGAGTCTTGTACACAGACGACCATGATGATGTTGGTGGTTCATATAGTCTCC 2172  
QY 1907 TCCTCAGCAGACATCTCTGGTATTCACTTCGGAACAGATGATGCGATCGGATCGAGCG 1966  
Db 2173 TCCTCAGTGCCTCTCTGGTATTCACTTCGGAACAGATGATGCGATCGGATCGAGCG 2232  
QY 1967 CTTGTTTACAGCAGACCTCTGGTATTCACTTCGGAACAGATGATGCGATCGGATCGAGCG 2026  
Db 2233 CTTGTGTCAGCAGACCTCTGGTATTCACTTCGGAACAGATGATGCGATCGGATCGAGCG 2292  
QY 2027 TCTCGTGGCGCTGGCACTGATGACACAGAGAAAGTTAAATCAGATGTTTTCCTCA 2086  
Db 2293 CTTCTGGAGTGGCACTGATGACACAGAGAAAGTTAAATCAGATGTTTTCCTCA 2352  
QY 2087 AAAGAACTCTTGACCATGACAGATGTCACAGACGACACAGATGTTTACAGCTGCACGCA 2146  
Db 2353 AAAGAACTCTTGACCATGACAGATGTCACAGACGACACAGATGTTTACAGCTGCACGCA 2412  
QY 2147 ACACCA 2153  
Db 2413 ATACCA 2419

RESULT 6  
US-09-245-041-12  
; Sequence 12, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq For Windows Version 3.0  
; SEQ ID NO 12  
; LENGTH: 6370  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-245-041-12

Query Match 35.7%; Score 1364.4; DB 3; Length 6370;  
Best Local Similarity 95.0%; Pred. No. 0;  
Matches 1492; Conservative 7; Mismatches 48; Indels 23; Gaps 8;

QY 2257 ATGTACTACTGTACAAAGACGACGTCGAGGAGCTGCGCTGGACCACTGCGAG 2316  
Db 1 ATGTACTACTGTACAAAGACGACGTCGAGGAGCTGCGCTGGACCACTGCGAG 60  
QY 2317 TGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCGAAATATCTGTGCAATTGGCTGG 2376  
Db 61 TGGAGCCCCGGAATCAGGAGTGCATTGCCCTGCCGAAATATCTGTGCAATTGGCTGG 120  
QY 2377 CATTTGGTGAAGACTCATGTTTGAATTTACTACTGCGAGGAGATTTATGCAATGCT 2436  
Db 121 CATTTGGTGAAGACTCATGTTTGAATTTACTACTGCGAGGAGATTTATGCAATGCT 180  
QY 2437 AAATTTGTTCTAGGAACCAATGCTTTCCTTACCAACCCAGAGAGGTTA 2496  
Db 181 AAATTTGTTCTAGGAACCAATGCTTTCCTTACCAACCCAGAGAGGTTA 240  
QY 2497 GAATTTGCTTAAGCAGCTGCGCAATATGCAATGTCATCTCAGAGCATGTCACAGCTCAC 2556  
Db

Db 241 GAATTTGTCTTAAGCAGCTGCGCAATATGCAATGTCATCTCAGAGCATGTCCAAAGCTCAC 300  
QY 2557 TTAACCCCATGGTTCGGCCCTT---CGGAAGATCAATGTGTCTTACTGG-----TGCTGGG 2608  
Db 301 TTAACCCCATGGTTCGGCCCTTTCGGGAAGGTCAATGTGTCTTACTKGTGTCTGGGAAG 360  
QY 2609 AAGATATGTCCCATTTTACAAATAG-TTTACTACAGTGGATGCCGTCTG-----AGCCAG 2663  
Db 361 GATATGTCTCCCATTTTACAAATAGTTTACTACAGTGGATGCCGTCTCTTGAAGCCAG 420  
QY 2664 TGATGCT--GGATTTCTGTGGAATTTT-----ATCAGAACCCAGTACTCGGGGACTCAAGG 2716  
Db 421 TGTGTCTTGGATTTCTGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGGACTCAAGG 480  
QY 2717 CTGCAACCTGAT--CAACCCACTCAATGTGTCTGTGAAAGCCCTGCACAAACACAG 2774  
Db 481 CTGCAACCTGATTCACCCACTTMAATGTGTGTCTGTGAAAGCCCTGCACAAACACAG 540  
QY 2775 TGCTAA--GCAGTGCAGACACCATGTCCTTGAGGACAGCATGTGAGATTCGACACAG 2833  
Db 541 TGCTAAGCAGTGCAGACACCATGTCCTTGAGGACAGCATGTGAGATTCGACACAG 600  
QY 2834 GCAGCTTGAGTGCATGTGTGCAACACATGAAGCAGTGTGTGACTCCAATGCTCATG 2893  
Db 601 GCAGCTTGAGTGCATGTGTGCAACACATGAAGCAGTGTGTGACTCCAATGCTCATG 660  
QY 2894 TGGCCTCTTCCCTTTTGGCCAGTGTATGGAATGTTATAGGACAGCATGCCCCCTG 2953  
Db 661 TGGCCTCTTCCCTTTTGGCCAGTGTATGGAATGTTATAGGACAGCATGCCCCCTG 720  
QY 2954 AAAATTTGTTGAGGCTACTGTACCTGTAGTCAATGCTTGAGCAACACAGCTGTGGCTGT 3013  
Db 721 AAAATTTGTTGAGGCTACTGTACCTGTAGTCAATGCTTGAGCAACACAGCTGTGGCTGT 780  
QY 3014 GTACTGATCCAGCAATCTGCAAGGAAATGCAATGAGGTTCTTATAAGGACCAAG 3073  
Db 781 GTACTGATCCAGCAATCTGCAAGGAAATGCAATGAGGTTCTTATAAGGACCAAG 840  
QY 3074 TGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCAGACCCCTGCTCAATTCCA 3133  
Db 841 TGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCAGACCCCTGCTCAATTCCA 900  
QY 3134 GCATGTGTAGGACACAGATCAACTGTGCTTTTCAATTCATCTGTCAGTTCGCAAT 3193  
Db 901 GCATGTGTAGGACACAGATCAACTGTGCTTTTCAATTCATCTGTCAGTTCGCAAT 960  
QY 3194 GCACGCGCCACAGTAAATCATCATCATGAGCATCTGTGAGAGTGTGAGACCTGACCA 3253  
Db 961 GCACGCGCCACAGTAAATCATCATCATGAGCATCTGTGAGAGTGTGAGACCTGACCA 1020  
QY 3254 CAGGCAAGCAGCTGCGAGACCTGCATATCTGCTTCTAGCGTATCCCAATGAGGGA 3313  
Db 1021 CAGGCAAGCAGCTGCGAGACCTGCATATCTGCTTCTACGGTATCCCAATGAGGGA 1080  
QY 3314 AATGTCAGCCTGCAAGTGCATGGCAGCGGTCTCTGTGCAACACCAACAGCGGCAAGT 3373  
Db 1081 AATGTCAGCCTGCAAGTGCATGGCAGCGGTCTCTGTGCAACACCAACAGCGGCAAGT 1140  
QY 3374 GCTTCTGCAACCAAGCGGTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAATC 3433  
Db 1141 GCTTCTGCAACCAAGCGGTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAATC 1200  
QY 3434 GATACAGGAAACCTCTCAGAGAAATGTTATATATCTCTTCTTATGACTATCAGT 3493  
Db 1201 GATACAGGAAACCTCTCAGAGAAATGTTATATATCTCTTCTTATGACTATCAGT 1260  
QY 3494 TCACCTTTAGTCTATCCAGGAAGATGATCGTATTTATATATCTCTTCTTATGACTATCAGT 3553  
Db 1261 TCACCTTTAGTCTATCCAGGAAGATGATCGTATTTATATATCTCTTCTTATGACTATCAGT 1320  
QY 3554 CTCCTGACGAAACAAACAGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTTCAACC 3613  
Db 1321 CTCCTGACGAAACAAACAGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTTCAACC 1380



QY 3614 TCAACATACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGAGAGAGATCCCTG 3673  
DB 1381 TCAACATACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGAGAGAGATCCCTG 1440  
QY 3674 TTGTTTCAAAAACCAACATTAAAGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 3733  
DB 1441 TTGTTTCAAAAACCAACATTAAAGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATT 1500  
QY 3734 TTGCAACCAACCAATATACATTTCTTTTATGTCAGTAATTTCACTGGCCCCATCA 3793  
DB 1501 TTGCAACCAACCAATATACATTTCTTTTATGTCAGTAATTTCACTGGCCCCATCA 1560  
QY 3794 AAATTCAGGT 3803  
DB 1561 AAATTCAGAT 1570

RESULT 7  
US-09-245-041-10  
; Sequence 10, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1998-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 10  
; LENGTH: 1051  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-10

Query Match 16.8%; Score 640.2; DB 3; Length 1051;  
Best Local Similarity 86.1%; Pred. No. 7.3e-155;  
Matches 724; Conservative 0; Mismatches 108; Indels 9; Gaps 1;  
QY 3 GGTGCGCCGAGGAGCGCGCGGCACTGAGGCAAGGCTGAGGAGGAGGAGCGCGGCGGAGC 62  
DB 216 GGTGCGCGCGCGCGGCGGCACTGAGGCGCGCTGAGGCGGAGCGAGCGAGCGAGCGC 275  
QY 63 GCTCGCGGCGAGGAGCGCGCGGCGCACTGGGACTGGGACTGGACCTGACCGGCTGGGAGGCC 122  
DB 276 GCCTCGCGGCGAGGAGGCGGAGGAGCGAGCGAGCGAGCGAGCGAGCGAGCGCGC 335  
QY 123 GGGGCTGGGGCGCGGCTGGCGCTCCCGGCTGCTGCTCCAGCGCTGGCGGCGCGCGCCT 182  
DB 336 GGGGCGCGCGCGCGCTGCTGCTCCCGGCTGCTGCTCGCGGCGCTGCTCGCGCGCGC 395  
QY 183 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 242  
DB 396 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 446  
QY 243 CGAGCGCGCGCGCGCGCGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 302  
DB 447 CGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 506  
QY 303 TCACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 362  
DB 507 TGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 566  
QY 363 CGCCCGCGCTGGGCGAGCAATGCGAGCACTGCGGGGCGCGCTTTCAGACTAACTGGATC 422  
DB 567 CACGGGCTGGGCGAGCAATGCGAGCACTGCGGGGCGCGCTTTCAGACTAACTGGGCTC 626

QY 423 TTCTGGGTTTGTGACAGATGACCTGGAAATATATAATACAAAACGAAGTGCACGTGCT 482  
DB 627 TTCTGGATTTGTAACAGATGACCTGGAAATATATAATATAAGACGAAGTGCACATGCT 686  
QY 483 CATTGAAGGACAGCGCAAAATAGAATAATGAGACTTCGTTTCAATCATTTTCTACAGATG 542  
DB 687 CATTGAAGGACAGCGCAAAATAGAATAATGAGACTTCGTTTCAACCAATTTTGTACAGAATG 746  
QY 543 TAGTTGGGACCAATTATATGTTTATGATGGGAGCTCAATTTATGACCGCTAGTTGCTGTC 602  
DB 747 TAGCTGGGACCAATTATATGTTTATGATGGGAGCTCAATCTACGCACTCTGATTTGCTGC 806  
QY 603 ATTTAGTGGCTCAATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCTTGAGGTTGTTGC 662  
DB 807 CTTTATGTGGCTCATTTGTTCTGAAAGAGATGGCAATGAGACGCGCTCCTGAGTCACTGT 866  
QY 663 CACATCAGTTATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 722  
DB 867 CACTTCAGTTATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 926  
QY 723 TAATATTACTTACAGTTTGTATGTTTGTCCAAATAACTGCTCAGCCGAGGAGAGTGAA 782  
DB 927 TAATATCACTTACAATTTGACATGCTGCGGAATAATTTGCTCAGCGCGAGGAGAGTGAA 986  
QY 783 GATCAGTATAGCAGCGAAACTGTTGAATGTAATGTTCTGNAAACTGGAAGGTGAAGC 842  
DB 987 GAGCAGTAACAGCAGCGGCTGTTGAGTGTGAATGTTCTGAAAACTGAAAAAGGGCGCG 1046  
QY 843 A 843  
DB 1047 A 1047

RESULT 8  
US-09-702-705-966/c  
; Sequence 966, Application US/09702705  
; Patent No. 6504010  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedvick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C14  
; CURRENT APPLICATION NUMBER: US/09/702,705  
; CURRENT FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 1833  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 966  
; LENGTH: 246  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-702-705-966

Query Match 6.4%; Score 246; DB 4; Length 246;  
Best Local Similarity 100.0%; Pred. No. 1e-53;  
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2515 CTGCGAATAATGAGTCACTCAGAGCATCTCAGAGCATGTCGAAGTCACTTAACCCCATGGTCCG 2574  
DB 246 CTGCGAATAATGAGTCACTCAGAGCATGTCGAAGTCACTTAACCCCATGGTCCG 187  
QY 2575 CTTGCGAAGTCAATGTTCTTCTGCTGCGAAGATATGTCCTTCAAAATAGT 2634  
DB 186 CTTGCGAAGTCAATGTTCTTCTGCTGCGAAGATATGTCCTTCAAAATAGT 127

```
QY 2635 TTACTACAGTGGAGCCGCTCTGAGCCAGTGGATGCTGGATTTGTGGAAATTTTATCAGAA 2694
Db 126 TTACTACAGTGGAGCCGCTCTGAGCCAGTGGATGCTGGATTTGTGGAAATTTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAGCTCAACCTGCATCAACCCCACTCAATGGTAGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGCTCAACCTGCATCAACCCCACTCAATGGTAGTCTGT 7
QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 9
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-966

Query Match 6.4%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 1e-53;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCATCTCAGAGCATGTCGAAGCTCAACCTCAACCCCACTGGTGGC 2574
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCGAAGCTCAACCTCAACCCCACTGGTGGC 187
QY 2575 CTTCCGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAATAAGT 2634
Db 186 CTTCCGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAATAAGT 127
QY 2635 TTACTACAGTGGAGCCGCTCTGAGCCAGTGGATGCTGGAAATTTTATCAGAA 2694
Db 126 TTACTACAGTGGAGCCGCTCTGAGCCAGTGGATGCTGGAAATTTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAGCTCAACCTGCATCAACCCCACTCAATGGTAGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGCTCAACCTGCATCAACCCCACTCAATGGTAGTCTGT 7
QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 10
US-09-614-124B-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
```

```
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-966

Query Match 6.4%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 1e-53;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2515 CTGCGAATAATGCAGTCATCTCAGAGCATGTCGAAGCTCAACCTCAACCCCACTGGTGGC 2574
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCGAAGCTCAACCTCAACCCCACTGGTGGC 187
QY 2575 CTTCCGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAATAAGT 2634
Db 186 CTTCCGAAGATCAATGTCTCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAATAAGT 127
QY 2635 TTACTACAGTGGAGCCGCTCTGAGCCAGTGGATGCTGGAAATTTTATCAGAA 2694
Db 126 TTACTACAGTGGAGCCGCTCTGAGCCAGTGGATGCTGGAAATTTTATCAGAA 67
QY 2695 CCCAGTACTCGGGGACTGAAGCTCAACCTGCATCAACCCCACTCAATGGTAGTCTGT 2754
Db 66 CCCAGTACTCGGGGACTGAAGCTCAACCTGCATCAACCCCACTCAATGGTAGTCTGT 7
QY 2755 GAAAGG 2760
Db 6 GAAAGG 1

RESULT 11
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-966

Query Match 6.4%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 1e-53;
```

	Matches	246;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;		
QY	2515	CTGCGAAT	TAATG	CAGTCA	TCTCTG	AGAGCAT	GTGCCA	AGCTCA	CTTAACCC	CATGGT	CGGC	2574
Db	246	CTGCGAAT	TAATG	CAGTCA	TCTCTG	AGAGCAT	GTGCCA	AGCTCA	CTTAACCC	CATGGT	CGGC	187
QY	2575	CTTCGGA	AGATCA	ATGTG	TCTACT	CTGGT	CTGCGA	AGATAT	GTCCCA	CTTTACAA	AATAGT	2634
Db	186	CTTCGGA	AGATCA	ATGTG	TCTACT	CTGGT	CTGCGA	AGATAT	GTCCCA	CTTTACAA	AATAGT	127
QY	2635	TTACTAC	AGTGGAT	CGCGT	CTGAG	CCCGAT	GTGCTG	GGATCT	GTGGAA	TTTAT	TATCAGAA	2694
Db	126	TTACTAC	AGTGGAT	CGCGT	CTGAG	CCCGAT	GTGCTG	GGATCT	GTGGAA	TTTAT	TATCAGAA	67
QY	2695	CCCACTACT	CGGGACT	GGAAG	GGCTG	CAAC	CTGCAT	CAACCC	CACTCA	ATGGTAGT	GTCTGT	2754
Db	66	CCCACTACT	CGGGACT	GGAAG	GGCTG	CAAC	CTGCAT	CAACCC	CACTCA	ATGGTAGT	GTCTGT	7
QY	2755	GAAAGG	2760									
Db	6	GAAAGG	1									

RESULT 12

US-09-833-381-1916

Sequence 1916, Application US/09833381

Patent No. 8672186

GENERAL INFORMATION:

APPLICANT: Robison, Keith E.

TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs

FILE REFERENCE: 5800-119

CURRENT APPLICATION NUMBER: US/09/833,381

CURRENT FILING DATE: 2001-04-11

PRIOR FILING DATE: 2000-02-29

NUMBER OF SEQ ID NOS: 2050

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 1916

LENGTH: 415

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(415)

OTHER INFORMATION: n = A,T,C or G

US-09-833-381-1916

	Query Match	5.6%	Score 212.8;	DB 4;	Length 415;
	Best Local Similarity	73.6%;	Pred. No. 4.9e-45;		
	Matches 271;	Conservative 0;	Mismatches 97;	Indels 0;	Gaps 0;
QY	422	CTTCCTGGTTGTGACATGGACCTCGAAATTATAAATACAAACGAAGTCACGTGGC	481		
Db	5	CGTCGGATATTTACAGATGGCCCAATTAACTATAATATAAACTAAATGTACTTGGC	64		
QY	482	TCATTGAAGGACAGCCAAATAGATTAATGAGACTTCGTTTCAATCTTTGCTACAGAGT	541		
Db	65	TCATTGAAGGGTATCCAAATGCAGTGTTTAAGATTAAAGATTCAATCAATTTGCTACAGAAT	124		
QY	542	GTAGTTGGACCATTTATGTTTATGATGGGGACTCAATTTATGCACCGGTAGTTGCTG	601		
Db	125	GTAGCTGGGATCATATGATGTTTATGATGGAGATTCAATATATGCACCTTAATAGCTG	184		
QY	602	CATTTAGTGGCTCATTTGTTCCGTGAGAGATGGCAATGAGACTGTCCTGAGGTTGTTG	661		
Db	185	TACTTAGTGGTTGTAGTCCCTGAAATAAGGGGCAATGAAACTGTGCTGAAGTTGTTA	244		
QY	662	CCACATCAGGTTATGCCCTTGTGTCATTTTTTTTAGTGTATGCTCTCTTAATATTGACTGGAT	721		
Db	245	CTACATCTGGCTATGCACGTGTTACATTTTTTTTAGTGTATGCTGCGTATAATCTAACTGGTT	304		
QY	722	TTAATATTACTTACAGTTTGTATATGTGCCAAATAAAGTCTCAGCGCAGAGAGTGTGA	781		

DB 305 TCACATTTTCTATTCAATCAATCTTGTCCTTAACAATTTGCTCTGGTCAIGGGAAGGAAAGIA 384

QY 782 AGATCAGT 789  
|  
|  
|  
|

Db 365 CAACTAGT 372

RESULT 13  
US-09-245-041-5  
; Sequence 5, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 90050  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-5

Query Match 5.3%; Score 201.2; DB 3; Length 90050;  
Best Local Similarity 81.5%; Pred. No. 1.le-40;  
Matches 233; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1169 TTGTTAGATATGTCTATTCTTTGGCATTTATACAAGATAAAAATTTACATGTATGGAGGAA 1228  
Db 6931 TGGTTTTAAANCTCGTTTACATCTTTCTTAGGATAAATCTACATGTATGGAGGAA 6990

QY 1229 AAATGATCCAACTGGGAATGTACCNAATGAGTTGAGAGTTTTTTCACATTCATAATGAGT 1288  
Db 6991 AAATGATTCACAGGGAACGTGCCAATGAGCTGAGAGTATTTTATTTCAATATGAT 7050

QY 1289 CATGGGTGTTGTGCACCCCTAACGCCAAAGAGCAGATGCGAGTGGTGGGCACCTCTGCAC 1348  
Db 7051 CATGGGTATTGTTAACTCCGAAAGCTAAGGATCAGTATGCGAGTTGGACACTCAGCAC 7110

QY 1349 ACATTTGTTACATGAAGAATGGCCGAGTGTGTATGCTGTCATCTTTGTCACCTGCCCTC 1408  
Db 7111 ACATTTGTTACACTGGCATCTGGCCGTGTGCTCATTTGGTTCATCTTGGTCATTTGCCAC 7170

QY 1409 TCTATGGATATAAGCAATGTCAGGAATATGATTTGGATAAGAA 1454  
Db 7171 TCTATGGATATAAGCGTTGTGCGAGGATATGACTTTGGGTATGTA 7216

RESULT 14  
US-09-245-041-3/c  
; Sequence 3, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 1, 2004, 07:06:46 ; Search time 6893 Seconds  
(without alignments)  
6190.781 Million cell updates/sec

Title: US-09-787-097-12  
Perfect score: 1429  
Sequence: 1 MVAATAATEARLRRTAATA.....KSGAVNRKQPPAQPGTCI 1429

Scoring table:  
OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 1

Total number of hits satisfying chosen parameters: 55025477

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO\_spool\_P/US09787097/runat\_26022004\_063015\_23757/app\_query.fasta\_1.1607  
-DB=EST -QFMT=fastap -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100  
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL -OUTFMT=pto  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09787097 @CGN 1.1 4337 runat\_26022004\_063015\_23757 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :  
EST:\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_nam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	805	56.3	3976	29	AY418587	AY418587 Homo sapi
2	252	17.6	1201	13	EX440935	EX440935 EX440935
3	227	15.9	681	12	BM783739	BM783739 K-EST0061
4	226	15.8	680	12	BM783358	BM783358 K-EST0061
5	226	15.8	680	12	BM783773	BM783773 K-EST0061
6	225	15.7	677	12	BM783718	BM783718 K-EST0061
7	224	15.7	673	12	BM783685	BM783685 K-EST0061
8	210	14.7	632	12	BM773003	BM773003 K-EST0057
9	210	14.7	632	12	BM783264	BM783264 K-EST0061
10	206	14.4	743	14	CB519923	CB519923 UI-M-GIO-
11	205	14.4	3976	29	AY418589	AY418589 Mus muscu
12	193	13.5	633	10	BE894518	BE894518 601431326
13	193	13.5	933	12	BG578679	BG578679 602824477
14	192	13.4	677	12	BM783821	BM783821 K-EST0061
15	185	12.9	917	10	BF672370	BF672370 602150605
16	171	12.0	683	12	BM783788	BM783788 K-EST0061
17	161	11.3	486	12	BM772968	BM772968 K-EST0057
18	155	10.8	962	13	BX431593	BX431593 BX431593
19	154	10.8	744	13	BQ183826	BQ183826 UI-H-EUO-
20	150	10.5	544	10	BF853962	BF853962 MR2-EN009
21	142	9.9	746	14	CB242644	CB242644 UI-CF-FNO
22	141	9.9	591	12	BI850289	BI850289 imageqc_1
23	136	9.5	643	14	CA306924	CA306924 UI-H-FTI-
24	136	9.5	991	10	BE880905	BE880905 601490448
25	135	9.4	3671	29	AY418588	AY418588 Pan trogl
26	134	9.4	560	10	BF853965	BF853965 MR2-EN009
27	134	9.4	782	14	CB521623	CB521623 UI-M-GHO-
28	133	9.3	598	10	BF853006	BF853006 MR2-EN009
29	129	9.0	739	14	CA502816	CA502816 UI-CF-FNO
30	122	8.5	503	14	CA895878	CA895878 B0195B01-
31	121	8.5	685	14	CA320165	CA320165 UI-M-FWO-
32	119	8.3	706	12	BI089332	BI089332 602853292
33	119	8.3	732	14	CD644584	CD644584 AGENCOURT
34	117	8.2	390	12	BG573763	BG573763 9010 cilia
35	116	8.1	493	10	BE259981	BE259981 601148539
36	115	8.0	399	10	BF880955	BF880955 QV1-ET018
37	115	8.0	681	12	BM981492	BM981492 UI-CF-EN1
38	113	7.9	674	12	BG563729	BG563729 602584512
39	111	7.8	400	9	AA350293	AA350293 EST57512
40	110	7.7	671	9	AI818764	AI818764 Wille04.X
41	108	7.6	502	9	AL048842	AL048842 DRFP2434G
42	104	7.3	484	10	BF963698	BF963698 PM4-NN120
43	99	6.9	383	9	AI372809	AI372809 EST175353
44	99	6.9	567	14	CA874419	CA874419 K0932C03-
45	99	6.9	953	10	BF313142	BF313142 601897661
46	98	6.9	595	10	BF853681	BF853681 MR2-EN009
47	95	6.6	565	14	CB525840	CB525840 UI-M-FYO-
48	93	6.5	320	9	AI372810	AI372810 EST175354
49	92	6.4	416	14	CB767839	CB767839 AMGNNUC:T
50	89	6.2	281	10	BG015062	BG015062 QV2-GN020
51	89	6.2	483	10	BF853021	BF853021 MR2-EN009
52	86	6.0	396	9	AA350292	AA350292 EST57511
53	84	5.9	758	14	CB522182	CB522182 UI-M-GHO-
54	84	5.9	778	13	BUS31905	BUS31905 AGENCOURT
55	82	5.7	484	12	BI004303	BI004303 PMO-HN007
56	81	5.7	840	14	CB520374	CB520374 UI-M-GIO-
57	80	5.6	416	14	R87660	R87660 Yp8901.s1
58	80	5.6	567	10	BF854487	BF854487 MR2-EN009
59	79	5.5	427	10	AW579586	AW579586 ME0-HT020
60	76	5.3	434	12	BI020540	BI020540 UI-M-CGDP
61	76	5.3	435	10	AW120845	AW120845 UI-M-BH2.
62	76	5.3	435	12	BI020684	BI020684 UI-M-CGDP

FEATURES		Location/Qualifiers	
source		1..3976	
gene		/organism="Homo sapiens"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9606"	
		<1..3976	
		/genes="ATRN"	
		/locus_tag="HCM6607"	
ORIGIN			
Alignment Scores:		Length: 3976	
Pred. No.:		Matches: 805	
Score:		Conservative: 0	
Percent Similarity:		100.00%	
Best Local Similarity:		100.00%	
Query Match:		Indels: 0	
DB:		Gaps: 0	
US-09-787-097-12 (1-1429) x AY418587 (1-3976)			
Qy	625	ThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThr	644
Db	1559	ACCATGATGTGTTCGGTGGTTTCAATAGTCTCTCTCAGCAGCATCTGGTATTCACC	1618
Qy	645	SerGluGlnCysAspAlaHisArgSerGluAlaAaCysLeuAlaGlyProGlyIle	664
Db	1619	TCGGAAACAGTGTGATGCGCATCGAGTGAAGCGCTGTGTAGCAGCAGGACCTGGTATT	1678
Qy	665	ArgCysValTrpAsnThrGlySerSerGlnCysIleSerTyrAlaLeuAlaThrAspGlu	684
Db	1679	CGTGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAA	1738
Qy	685	GlnGluGlnCysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCys	704
Db	1739	CAAGAGAAAGTAAATCAGATGCTTTTCCAAAAGAACCTTGACCATGACAGATGT	1798
Qy	705	AspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsn	724
Db	1799	GACCAGCACACAGATTGTTACAGTCGACAGCCACACCAATGACTGCCACTGGTGCAT	1858
Qy	725	AspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyr	744
Db	1859	GACCATTTGTGTCCTCCAGGAACACACAGTCGTCTAGAGGCCAGATCTCCATTTTAGTAT	1918
Qy	745	GluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSer	764
Db	1919	GAGATTTGCCCAAGGATAACCCCATGTACTACTGTAAACAAGAACACAGCTCGAGGAGC	1978
Qy	765	CysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuPro	784
Db	1979	TGTGCCCTGACCAAGAACCTGCCAGTGGAGCCCGGAATCAGAGATGTCATTCCTGCC	2038
Qy	785	GluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThr	804
Db	2039	GAATAATCTGTGGCATTTGGCTGGCATTTGGTTGGAACTCATGTTTGAATACTACT	2098
Qy	805	AlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuAla	824
Db	2099	GCCAAAGAGAAATATGACAAATGCTAAATTTGTCTGTAGGAACCAACAATCCCTTTGGT	2158
Qy	825	SerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSer	844
Db	2159	TCCTTTACACCCAGGAAGAGGTAGATTTGCTTTAACAGCTCGAATAATGCAATCA	2218
Qy	845	SerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnVal	864
Db	2219	TCTCAGAGATGTCTCAAGCTCACCTTAACCCATGGTGGCTTCGGAGATCAATGTG	2278
Qy	865	SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuTrpMetPro	884
Db	2279	TCTTACTGTGTGGAGAGATATGTCCTTACCAATAGTTTACTACAGTGGATGCCG	2338
Qy	885	SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu	904

2339	Db	TCTGAGCCCGAGTGGATCTCTGGAAATTTTATCAGAACCCAGTACTCGGGAGCTG	2398
905	QY	LYEAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProIleAsnHis	924
2399	Db	AAGGCTGCACCTGGCATCAACCCACATCAATGGTAGTGTCTGTGAAGAAGGCTGCNAACCAAC	2458
925	QY	SerAlaIysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSer	944
2459	Db	AGTGTAAAGCAGTGGCGGACCACTGTCCTTGAGGACAGCATGTGGAGATTGCACACAGC	2518
945	QY	GlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyr	964
2519	Db	GGCAGCTCTGAGTGCATGTGGTGCAGCAATGACGAGTGTGTGACTTCCAATGGCTAT	2578
965	QY	ValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProPro	984
2579	Db	GTGGCTCTCTCCCTTTGGCCAGTGTATGGAAATGGTATACGATGAGCACCTGCCCCCT	2638
985	QY	GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp	1004
2639	Db	GAAAAATGTTTCAGGCTACTGTACTGTAGTCATTGTTGGAGCAACGAGCTGTGGCTGG	2698
1005	QY	CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrIysGlyPro	1024
2699	Db	TGTACTGATCCAGCAATACTGGCAAAAGGAAATGCATAGAGGGTTCTTATAAGGACCA	2758
1025	QY	ValIysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer	1044
2759	Db	GTGAAGATGCTTCGGAAGCCCTACAGGAAATTTCTATCCAGGCCCCGTGTCAATCC	2818
1045	QY	SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProIlaCysGln	1064
2819	Db	AGCATGTGCTAGAGACAGCATCAACTGGTCTTCACTTCACTGTCCAGCTTGCCAA	2878
1065	QY	CysAsnGlyHisSerIysCysIleAsnGlnSerIleCysGluIysCysGluAsnLeuThr	1084
2879	Db	TGCAACGGCCACAGTAAATGCATCAATCAGAGATCTGTGAGAAGTGTGAGAACCTGACC	2938
1085	QY	ThrGlyIysHisCysGluThrCysIleSerGlyPheTyrGluVAspProThrAsnGlyGly	1104
2939	Db	ACAGGCAAGCAGCTGCGAGACTGTCATATCTGGCTTCTACCGTGTATCCCAATGGAGGG	2998
1105	QY	LysCysGlnProCysLysCysAsnGlyHisIleAlaSerLeuCysAsnThrAsnThrGlyLys	1124
2999	Db	AAATGTGAGCCATGCAAGTGCATGGCCAGCGTCTCTGTGCAACCAACACGGGCAAG	3058
1125	QY	CysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsn	1144
3059	Db	TGCTTCTGCACACCAAGGGCGTCAAGGGGAGCAGTGCCAGCTATGTGAGGTAGAAAAAT	3118
1145	QY	ArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGln	1164
3119	Db	CGATACCAAGAAACCTCTCAGAGGAACATGTTATTATATCTCTTATTGACTATCAG	3178
1165	QY	PheThrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIleAsnPheValAla	1184
3179	Db	TTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGGCT	3238
1185	QY	ThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsn	1204
3239	Db	ACTCTGACGAAACAACACAGGATTTGGACATGTTTCATCATGTGCTCCAGAATTTCAAC	3298
1205	QY	LeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetPro	1224
3299	Db	CTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCCT	3358
1225	QY	ValValSerLysThrAsnIleLysGluTyrIysAspSerPheSerAsnGluLysPheAsp	1244
3359	Db	GTTGTTTCAAAAAACAACATTAAAGAGTACAAAGATAGTTTCTCTTAATGAGAGTTTGAT	3418
1245	QY	PheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIle	1264

/dev stage="fetal"  
 /clone lib="Homo sapiens FETAL BRAIN"  
 /notes="Organ: brain; Vector: pCMVSPORT\_6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,77e-193 Length: 1201  
 Score: 252.00 Matches: 301  
 Percent Similarity: 98.69% Conservative: 0  
 Best Local Similarity: 98.69% Mismatches: 2  
 Query Match: 17.63% Indels: 4  
 DB: 13 Gaps: 0

US-09-787-097-12 (1-1429) x BX440935 (1-1201)

QY 780 CysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCys 799  
 DB 58 TGCAITGCCCTGCCGAAATATCTGTGCATGTGCGCATTTGGTTGGAACTCATGT 117  
 QY 800 LeuLysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHis 819  
 DB 118 TTGAAATTTACTACTGCCAAGGAGAAATTATGACAATGCTAAATTTCTGTAGGAACAC 177  
 QY 820 AsnAlaLeuAlaSerLeuThrThrGlnLys--LysValGluPheValLeuLysGlnL 839  
 DB 178 ATGCGCTTTTGGCTCTCTTACACCCCT-GA-TGAAGTAGAATTTGCTTTAGCAGC 235  
 QY 839 euArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyL 859  
 DB 236 TCGAATAATGAGTCATCTCAGAGCATGCTCAAGCTCACCTTAACCCCATGGGTGGCC 295  
 QY 859 euArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerL 879  
 DB 296 TTCGAAGATCAATGTCTACTGTCTGTGGTGGAAAGATATGTCCCATTTTCAAAATAGT 355  
 QY 879 euLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluP 899  
 DB 356 TACTACAGTGGATCGCTCTGAGCCAGTCATGTCTGGATTTCTGGAAATTTTATCAGAAC 415  
 QY 899 roSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysG 919  
 DB 416 CCAGTACTCGGGAGTGAAGGCTGCAACCTGCATCAACCCCATCAATGGTAGTGTCTGTG 475  
 QY 919 luArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaC 939  
 DB 476 AAAGGCTGCACCAACACAGTGTAGACAGTGCAGGACACCATGTGCTTGGAGCAGCAT 535  
 QY 939 ysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysV 959  
 DB 536 GTGGAGATTGCCACAGCGGAGCTCTGAGTGCATGTGTGTGTCAGCAACATGAAGCAGTGTG 595  
 QY 959 alAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrM 979  
 DB 596 TGGATCTCCATGCCATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTATACGA 655  
 QY 979 etSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluG 999  
 DB 656 TGAGCACCTGCCCCCTGAAATTTGTTCCAGGCTACTGTACCTGTAGTCAATGCTTGGAGC 715  
 QY 999 InProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluG 1019  
 DB 716 AACCAAGCTGTGGCTGT 775  
 QY 1019 lySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProG 1039  
 DB 776 GTTCTATAAAGGACCAAGTGAATGCTTCCGAAAGCCCTTACAGGAAATTTCTATCCAC 835  
 QY 1039 InProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheLeH 1059

DB 836 AGCCCTGCTCAATTCAGCATGTGTAGAGACAGCAGATACAACTGGTCTTTTCATTC 895  
 QY 1059 isCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluL 1079  
 DB 896 ACTGTCCAGCTTGCCATGCAAGCGGCACAGTAATGTCATCAATCAGAGCATCTGTGAGA 955  
 QY 1079 yscysgluasn 1082  
 DB 956 AGTGTGAGAAC 966

## RESULT 3

BM783739

## LOCUS

DEFINITION K-EST0061730 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-H05 5',  
 mRNA sequence.

## ACCESSION

BM783739

## VERSION

BM783739.1

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

## REFERENCE

1 (bases 1 to 681)

## AUTHORS

Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,  
 Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and  
 Kim, Y.S.

## TITLE

21C Frontier Korean EST Project 2001

## JOURNAL

Unpublished (2002)

## COMMENT

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 34 row: H column: 05

High quality sequence stop: 681.

## FEATURES

Location/Qualifiers

1..681

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S5SNU484-34-H05"

/sex="M"

/tissue\_type="Stomach"

/cell\_type="Epithelial"

/cell\_line="SNU-484"

/lab\_host="Top10F"

/clone\_lib="S5SNU484"

/notes="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;  
 Site 2: NotI; The poly (A) + RNA was decapped with tobacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 RNA ligase. The  
 first strand cDNA was synthesized from oligo dt-selected  
 mRNA by priming with dt-tailed vector. The dt-tailed  
 vector was adjusted to have about 60nt. The cDNA vector  
 was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a DNA strand by Okayama-Berg method.  
 The obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F, by electroporation  
 method."

## ORIGIN

## Alignment Scores:

Pred. No.: 1.88e-173 Length: 681  
 Score: 227.00 Matches: 227  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 15.89% Indels: 0  
 DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM783739 (1-681)



Qy 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerIleuThrLeuThrProTrrpVal 857  
 Db 1 CAGCTGCGAATAATGCACTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCATGGGTC 60  
 Qy 858 GlyLeuArgIleAsnValSerTrrpCysThrGluAspMetSerProPheThrAsn 877  
 Db 61 GGCCTTCGGAAGATCAATGTCTCTACTGGTGGGAGATATGTCCCATTTTACAAAT 120  
 Qy 878 SerLeuLeuGlnTrrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897  
 Db 121 AGTTTACTACAGTGGATCCGCTCTGAGCCAGTCATGCTGATTCTGGGAATTTATCA 180  
 Qy 898 GluProSerThrArgGlyLeuLysAlaThrCysIleAsnProLeuAsnGlySerVal 917  
 Db 181 GAACCCAGTACTCGGGAGTCAAGCTGCAACCTGTCATCAACCCACCTCAATGGTAGTGC 240  
 Qy 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937  
 Db 241 TGTGNAAGCCTGCAACACACAGTCTTAAGCAGTCCGCGACACCATGTGCTTGAGGACA 300  
 Qy 938 AlaCysGlyAspCysThrSerGlySerGluCysMetTrrpCysSerAsnMetLysGln 957  
 Db 301 GCATGTGAGATTCACACGAGCGCAGCTCTGAGTGCATGTGGTGGCCTTGAGGACA 360  
 Qy 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrrpTyr 977  
 Db 361 TGTGTGGACTCCATGCTCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGATGGTAT 420  
 Qy 978 ThrMetSerThrCysProGluAsnCysSerGlyTrrpCysThrCysSerHisCysLeu 997  
 Db 421 ACGATGAGCAGCTGCCCCCTGAAATTTGTTCAAGCTACTGCTAGCTGATGCTGTTG 480  
 Qy 998 GluGlnProGlyCysGlyTrrpCysThrAspProSerAsnThrGlyLysGlyCysValle 1017  
 Db 481 GAGCAACAGGCTGTGGTGTGTACTGATCCCAATATCTGGCAAGGGAATGCATA 540  
 Qy 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037  
 Db 541 GAGGGTTCTTATAAGGACCACTGAGATGCTTTCGCAAGCCCTACAGGAAATTTCTAT 600  
 Qy 1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrrpSerPhe 1057  
 Db 601 CCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACCACTGGTCTTTC 660  
 Qy 1058 IleHisCysProAlaCysGln 1064  
 Db 661 ATTCACTGTCCAGCTTGCCAA 681

RESULT 4  
 LOCUS BM783258  
 DEFINITION K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5',  
 mRNA sequence.  
 ACCESSION BM783258  
 VERSION BM783258.1 GI:19131490  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 680)  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr  
 Plate: 31 row: H column: 08  
 High quality sequence stop: 680.  
 Location/Qualifiers  
 1. .680  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S5SNU484-31-H08"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Epithelial"  
 /cell\_line="SNU-484"  
 /lab\_host="Top10F"  
 /clone\_lib="S5SNU484"  
 /notes="Organ: Stomach; Vector: pTZ18RP; Site\_1: EcoRI;  
 Site\_2: NotI; The poly (A)+ RNA was decapped with tobacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 RNA ligase. The  
 first strand cDNA was synthesized from oligo dt-selected  
 mRNA by priming with dt-tailed vector. The dt-tailed  
 vector was adjusted to have about 60nt. The cDNA vector  
 was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a DNA strand by Okayama-Berg method.  
 The obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10F' by electroporation  
 method."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,22e-172 Length: 680  
 Score: 226.00 Matches: 226  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 15.82% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-787-097-12 (1-1429) x BM783258 (1-680)  
 Qy 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerIleuThrLeuThrProTrrpVal 857  
 Db 1 CAGCTGCGAATAATGCACTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCATGGGTC 60  
 Qy 858 GlyLeuArgIleAsnValSerTrrpCysThrGluAspMetSerProPheThrAsn 877  
 Db 61 GGCCTTCGGAAGATCAATGTCTCTACTGGTGGGAGATATGTCCCATTTTACAAAT 120  
 Qy 878 SerLeuLeuGlnTrrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897  
 Db 121 AGTTTACTACAGTGGATGCGCTCTGAGCCAGTGTGCTGGATCTGTGGAAATTTATCA 180  
 Qy 898 GluProSerThrArgGlyLeuLysAlaThrCysIleAsnProLeuAsnGlySerVal 917  
 Db 181 GAACCCAGTACTCGGGAGTCAAGCTGCAACCTGTCATCAACCCACCTCAATGGTAGTGC 240  
 Qy 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937  
 Db 241 TGTGNAAGCCTGCAACACACAGTCTTAAGCAGTCCGCGACACCATGTGCTTGAGGACA 300  
 Qy 938 AlaCysGlyAspCysThrSerGlySerGluCysMetTrrpCysSerAsnMetLysGln 957  
 Db 301 GCATGTGAGATTCACACGAGCGCAGCTCTGAGTGCATGTGGTGGCCTTGAGGACA 360  
 Qy 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrrpTyr 977  
 Db 361 TGTGTGGACTCCATGCTCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGATGGTAT 420  
 Qy 978 ThrMetSerThrCysProGluAsnCysSerGlyTrrpCysThrCysSerHisCysLeu 997  
 Db 421 ACGATGAGCAGCTGCCCCCTGAAATTTGTTCAAGCTACTGCTAGCTGATGCTGTTG 480  
 Qy 998 GluGlnProGlyCysGlyTrrpCysThrAspProSerAsnThrGlyLysGlyCysValle 1017

Db 481 GAGCACACAGGCTGTGGCTGGTGTACTGATCCAGCATACTGGCAGGAAATGCATA 540  
 QY 1018 GluGlySerTyrIysGlyProValIysMetProSerGlnAlaProThrGlyAsnPheTyr 1037  
 Db 541 GAGGGTTCCTATTAAGAGACAGTGAAGATGCGCTTCGACGCCCTACAGGAAATTCAT 600  
 QY 1038 ProGlnProLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057  
 Db 601 CCACAGCCCTCTGCTCAATTCAGCATGTCTAGAGCAGCAGATACACTGGTCTTTC 660  
 QY 1058 IleHisCysProAlaCys 1063  
 Db 661 ATTCAGTCCAGCTTGC 678

RESULT 5  
 BM783773  
 LOCUS  
 DEFINITION K-EST0061775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5',  
 mRNA sequence.  
 ACCESSION BM783773  
 VERSION BM783773.1 GI:19132005  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 680)  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS

Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoen-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 36 row: D column: 02  
 High quality sequence stop: 680.

## FEATURES

source  
 1..680  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S5SNU484-36-D02"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Epithelial"  
 /lab\_host="Top10F"  
 /clone\_lib="S5SNU484"  
 /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;  
 Site 2: NotI; The poly (A)+ RNA was decapped with tobacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 RNA ligase. The  
 first strand cDNA was synthesized from oligo dt-selected  
 mRNA by priming with dt-tailed vector. The dt-tailed  
 vector was adjusted to have about 60nt. The cDNA vector  
 was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a DNA strand by Okayama-Berg method.  
 The obtained cDNA vectors were used for transfection of  
 competent cells E. coli Top10F by electroporation  
 method."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1-22e-172 Length: 680  
 Score: 226.00 Matches: 226  
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 15.82% Indels: 0  
 DB: 12 Caps: 0

US-09-787-097-12 (1-1429) x BM783773 (1-680)

QY 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerIysLeuThrLeuThrProTrpVal 857  
 Db 1 CAGTCGGATATATGAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCATGGGTC 60  
 QY 858 GlyLeuArgIysIleAsnValSerTyrTyrCysTyrGluAspMetSerProPheThrAsn 877  
 Db 61 GGCCTTCGGAAGATCAATGTGCTCTACTGCTGGTGGGAAGATATGTCCCAATTTACAAAT 120  
 QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897  
 Db 121 AGTTTACTACAGTGGATGCGCTCTGAGCCAGTGTGGATTCGTGGAAATTTATCA 180  
 QY 898 GluProSerThrArgGlyLeuIysAlaIaThrCysIleAsnProLeuAsnGlySerVal 917  
 Db 181 GAACCCAGTACTCGGGACTGAAGCGCTGCAACCTGCATCAACCCACTCAATGTAGTGTCTC 240  
 QY 918 CysGluArgProAlaAsnHisSerAlaIysGlnCysArgThrProCysAlaLeuArgThr 937  
 Db 241 TGTGAAGGCTGCAACACACAGTGTAGCAGTGCAGCACCACCATGTGCTTGAGGACA 300  
 QY 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTyrCysSerAsnMetIysGln 957  
 Db 301 GCATGTGGAGATTGCACCGGCGCAGCTCTGAGTGCATGTGGTGACAGCAACATGAAGCAG 360  
 QY 958 CysValAspSerAsnAlaTyrValIaSerPheProPheGlyGlnCysMetGluTrpTyr 977  
 Db 361 TGTGTGAGCTCCAAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTAT 420  
 QY 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997  
 Db 421 ACATGAGCAGCTGCCCCCTGANAATTTGTCAGGCTACTGTACCTGTAGTATTGCTTG 480  
 QY 998 GluGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyIysGlyCysIle 1017  
 Db 481 GAGCAACACCGGCTGTGGCTGGTGTACTGTATCCAGCAATACTGGCAAGGAAATGCATA 540  
 QY 1018 GluGlySerTyrIysGlyProValIysMetProSerGlnAlaProThrGlyAsnPheTyr 1037  
 Db 541 GAGGGTTCCTATTAAGAGCAGTGAAGATGCGCTTCGACGCCCTACAGGAAATTTCTAT 600  
 QY 1038 ProGlnProLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057  
 Db 601 CCACAGCCCTCTGCTCAATTCAGCATGTCTAGAGCAGCAGATACAACTGGTCTTTC 660  
 QY 1058 IleHisCysProAlaCys 1063  
 Db 661 ATTCAGTCCAGCTTGC 678

## RESULT 6

BM783773  
 LOCUS  
 DEFINITION K-EST0061702 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-F01 5',  
 mRNA sequence.  
 ACCSSION BM783718  
 VERSION BM783718.1 GI:19131950  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 677)  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.  
 TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS

Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 34 row: F column: 01  
High quality sequence stop: 677.  
Location/Qualifiers

# FEATURES

1. 677  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S5SNU484-34-F01"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/cell\_lines="SNU-484"  
/lab\_host="Top10F"  
/clone\_lib="S5SNU484"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method."

## ORIGIN

Alignment Scores:  
Pred. No.: 7, 83e-172 Length: 677  
Score: 225.00 Matches: 225  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 15.75% Indels: 0  
DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM783718 (1-677)

QY 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857  
DB 1 CAGCTGCGAATAATGACGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGTGC 60  
QY 858 GlyLeuArgGlyIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877  
DB 61 GGCCCTCGAAGATCAATGTGTCTTACTGTGTCTGGGAGATATGTCTCCCATTTACAAAT 120  
QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897  
DB 121 AGTTTACTACAGTGGATGCGGCTGAGCCAGTCAGTGTGGATTCTGTGGAATTTATCA 180  
QY 898 GluProSerThrArgGlyLeuIlysAlaIaThrCysIleAsnProLeuAsnGlySerVal 917  
DB 181 GAACCCAGTACTCGGGGAGTGAAGCTGCAACCTGCATCAACCCCACTCAATGGTAGTGTCT 240  
QY 918 CysGluArgProAlaAsnHisSerAlaIysGlnCysArgThrProCysAlaLeuArgThr 937  
DB 241 TGTGAAGGCTGCAACACACAGTGTGAAGCAGTCGCGACACCATGTGCTTGGAGACA 300  
QY 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957  
DB 301 GCATGTGGAGATTGACACGAGCGAGCTCTGAGTGCATGTGTGTGAGCAACATGAAGCAG 360  
QY 958 CysValAspSerAsnAlaIaThrValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977  
DB 361 TGTGTGGACTCCAAATGCGCTTANGTGGCTTCTTCCCTTTTGGCCAGTGTATGGATGGTAT 420

QY 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997  
DB 421 ACAGATGAGCACCTGCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCATTCCTTG 480  
QY 998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017  
DB 481 GAGCAACACAGGCTGTGGTGTACTGATCCAGCAATACTGGCAAGGGAATCATCA 540  
QY 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037  
DB 541 GAGGGTTCTATTAAGGACGAGTGAAGATGCTTCCCAAGCCCTACAGGAATTTCTAT 600  
QY 1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057  
DB 601 CCACAGCCCTGCTCAATCCAGCATGTGTCTAGAGCAGCAGCAGATACAACTGGTCTTTC 660  
QY 1058 IleHisCysProAla 1062  
DB 661 ATTCATGTCACGCT 675

## RESULT 7

BM783685  
LOCUS 673 bp mRNA linear EST 05-MAR-2002  
DEFINITION K-EST0061659 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-B06 5', mRNA sequence.  
ACCESSION BM783685  
VERSION BM783685.1 GI:19131917  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 673)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 34 row: B column: 06  
High quality sequence stop: 673.  
Location/Qualifiers

## FEATURES

1. 673  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S5SNU484-34-B06"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/cell\_lines="SNU-484"  
/lab\_host="Top10F"  
/clone\_lib="S5SNU484"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method."

## ORIGIN

## Alignment Scores:

Pred. No.: 5,03e-171 Length: 673  
 Score: 224.00 Matches: 224  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 15.68% Indels: 0  
 DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM783685 (1-673)

QY 838 GlnLeuArgIleMetGlnSerGlnSerMetSerLysLeuThrLeuThrProTyrVal 857  
 DB 1 CAGCTGCGAATATGAGTCACTCTGAGCATGTCAGACATGCTCAAGCTCACCTTAACCCATGGGTC 60  
 QY 858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877  
 DB 61 GGCCTTCGGAAGATCAATGTCTACTGCTGCTGCGAAGATATGTCACCATTTACAAT 120  
 QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897  
 DB 121 AGTTTACTACATGATGCGCTGAGCCAGTATGCTGGATTCTGTGGAAATTTATCA 180  
 QY 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917  
 DB 181 GAACCCAGCTACTCGGGAGCTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTCTC 240  
 QY 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937  
 DB 241 TGTGAAGGCCCTGCAACACACAGTCTAAGCAGTCCGACACCATGTGCTTGGAGACA 300  
 QY 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957  
 DB 301 GCATGTGAGATTGACACAGCGGAGCTCTGATGCTGCTGGTGAGCAATCATGAGCAG 360  
 QY 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977  
 DB 361 TGTGTGGACTCCAAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTAT 420  
 QY 978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997  
 DB 421 ACATGAGACCTGCTGCCCTGCAAAATGTTGAGCTACTGTACTGTAGCTATGCTGCTG 480  
 QY 998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyIleCysCysIle 1017  
 DB 481 GAGCAACACAGGCTGTGGCTGTGTACTGATCCAGCAATCTGCAAGAGGAATGCATA 540  
 QY 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037  
 DB 541 GAGGGTTCCTATAAAGGACCATGAGATGCTTCCCAAGGCCCTTACAGGAATTTCTAT 600  
 QY 1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057  
 DB 601 CCACAGCCCTCTCAATTCAGCATGTGCTAGAGGACAGCAGATACAACTGGTCTTTC 660  
 QY 1058 IleHisCysPro 1061  
 DB 661 ATTCACTGTCCA 672

## RESULT 8

BM773003 632 bp mRNA linear EST 04-WAR-2002  
 LOCUS K-EST0057379 S5SNU484s1 Homo sapiens cDNA clone S5SNU484s1-3-E09  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BM773003  
 VERSION BM773003.1 GI:19102618  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 632)  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

## TITLE

JOURNAL

## COMMENT

Oh,K.J., Cheong,J.B., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.  
 21C Frontier Korean EST Project 2001  
 Unpublished (2002)  
 Contact: Kim YS  
 Genome Research Center  
 Korea Research Institute of Bioscience & Biotechnology  
 52 Eoem-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongsung@mail.kribb.re.kr  
 Plate: 3 row: E column: 09  
 High quality sequence stop: 632.

## FEATURES

## source

1..632  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S5SNU484s1-3-E09"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Epithelial"  
 /cell\_line="SNU-484"  
 /lab\_host="Top10F"  
 /clone\_lib="S5SNU484s1"  
 /notes="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA digestion was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F, by electroporation method. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and NidT14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the subtracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F, with electroporation method."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,06e-159 Length: 632  
 Score: 210.00 Matches: 210  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 14.70% Indels: 0  
 DB: 12 Gaps: 0  
 US-09-787-097-12 (1-1429) x BM773003 (1-632)  
 QY 838 GlnLeuArgIleMetGlnSerGlnSerMetSerLysLeuThrLeuThrProTyrVal 857  
 DB 1 CAGCTGCGAATATGAGTCACTCTGAGCATGTCAGACATGCTCAAGCTCACCTTAACCCATGGGTC 60  
 QY 858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877  
 DB 61 GGCCTTCGGAAGATCAATGTCTACTGCTGCGAAGATATGTCACCATTTACAAT 120  
 QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897  
 DB 121 AGTTTACTACATGATGCGCTGAGCCAGTATGCTGGATTCTGTGGAAATTTATCA 180

Qy 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917  
 Db 181 GAACCACTACTCGGGGACTGAGGCTGCACCTGCATCAACCCACTCAATGGTAGTGC 240  
 Qy 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937  
 Db 241 TGTGAAGGCTCTGCAACACACAGTCTAAGCAGTCCGGACACCATGTGCTTGAGGACA 300  
 Qy 938 AlaCysGlyAspCysThrSerGlySerGluCysMetTrpCysSerAsnMetLysGln 957  
 Db 301 GCATGTGGAGATTGCACACGCGGAGCTCTGAGTGCATGTGTGCGCAACATGAAGCAG 360  
 Qy 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977  
 Db 361 TGTGTGGACTCCAAATGCCATATGTGGCTCTCTTCCCTTTTGGCCAGTATGGAATGGTAT 420  
 Qy 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997  
 Db 421 ACATGAGCAGCTGCCCCCTGAAATTTGTTCCAGCTACTGTACCTGTAGTCATGCTGTG 480  
 Qy 998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017  
 Db 481 GAGCAACAGAGCTGTGGCTGTGTACTGATCCACCAATACTGGCAAGGGAATGCATA 540  
 Qy 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037  
 Db 541 GAGGGTTCTCTATAAAGGACCACTGAAGATGCCCTTCGCAAGCCCTTACAGGAATTTCTAT 600  
 Qy 1038 ProGlnProLeuLeuAsnSerMetCys 1047  
 Db 601 CCACAGCCCTCTCAATTCAGCATGTGT 630

## RESULT 9

BM783264

LOCUS

DEFINITION K-EST0061166 S55NU484 Homo sapiens cDNA clone S55NU484-32-A05 5',  
 mRNA sequence.

ACCESSION

BM783264

VERSION

BM783264.1

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1. (bases 1 to 632)

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.

21C Frontier Korean EST Project 2001

Unpublished (2002)

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience &amp; Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 32, row: A column: 05

High quality sequence stop: 632.

Location/Qualifiers

1. .632

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S55NU484-32-A05"

/sex="M"

/tissue\_type="Stomach"

/cell\_type="Epithelial"

/cell\_line="SNU-484"

/lab\_host="Top10F"

/clone\_lib="S55NU484"

/note="Organ: Stomach; Vector: pTZ18RP1; Site\_1: EcoRI;

FEATURES

source

## ORIGIN

Alignment Scores:

Pred. No.: 1,06e-159 Length: 632

Score: 210.00 Matches: 210

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 14.70% Indels: 0

DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM783264 (1-632)

Qy 838 GluLeuArgIleXetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTropVal 857  
 Db 1 CAGCTGGGAATAATGAGTCAATCTCAGAGCATGTCGAAGCTCACCTTAACCCCATGGGTC 60  
 Qy 858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877  
 Db 61 GGCCTTCGGAAGATCAATGTGTCTACTGTGTGGAGATATGTCCTCATTTACAAAT 120  
 Qy 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897  
 Db 121 AGTTTACTACAGTGGATGCCGTCTGAGCCAGTGTGCTGGATCTGTGGAATTTTATCA 180  
 Qy 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917  
 Db 181 GAACCCAGTACTCGGGAGCTGAAGGCTGCACTGCACTCAACCCACTCAATGGTAGTGC 240  
 Qy 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937  
 Db 241 TGTGAAGGCTCTGCAACACACAGTCTAAGCAGTCCGGACACCATGTGCTTGAGGACA 300  
 Qy 938 AlaCysGlyAspCysThrSerGlySerGluCysMetTrpCysSerAsnMetLysGln 957  
 Db 301 GCATGTGGAGATTGCACACGCGGAGCTCTGAGTGCATGTGTGCGACCAACATGAAGCAG 360  
 Qy 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977  
 Db 361 TGTGTGGACTCCAATGCCATGTGTGGCTCTCTTCCCTTTTGGCCAGTATGGAATGGTAT 420  
 Qy 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997  
 Db 421 ACATGAGCAGCTGCCCCCTGAAATTTGTTCCAGCTACTGTACCTGTAGTCATGCTGTG 480  
 Qy 998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017  
 Db 481 GAGCAACAGAGCTGTGGCTGTGTACTGATCCACCAATACTGGCAAGGGAATGCATA 540  
 Qy 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037  
 Db 541 GAGGGTTCTCTATAAAGGACCACTGAAGATGCCCTTCGCAAGCCCTTACAGGAATTTCTAT 600  
 Qy 1038 ProGlnProLeuLeuAsnSerMetCys 1047  
 Db 601 CCACAGCCCTCTCAATTCAGCATGTGT 630

## RESULT 10

CB519923

LOCUS

DEFINITION UI-M-G10-cek-g-05-0-UI\_r1 NIH BMAP\_G10 Mus musculus cDNA clone  
 IMAGE:6840702 5', mRNA sequence.

743 bp mRNA linear EST 09-JUL-2003

ACCESSION CB519923  
 VERSION 1  
 KEYWORDS GI:29353278  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 743)  
 REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at <http://genome.uiowa.edu/distribution/mousefl.html>  
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)  
 Seq primer: pYX-5.  
 Location/Qualifiers  
 1..743  
 /organism="Mus musculus"  
 /mol\_type="rRNA"  
 /db\_xref="taxon:10090"  
 /clone IMAGE:6840702  
 /tissue\_type="whole brain"  
 /dev\_stage="embryo 13.5,14.5,16.5,17.5dpc"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH\_BMAP\_G10"  
 /note="Organ: Brain; Vector: pYX-Asc; Site\_1: EcoR I; Site\_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 2,17e-156 Length: 743  
 Score: 206.00 Matches: 206  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 14.42% Indels: 0  
 DB: 14 Gaps: 0  
 US-09-787-097-12 (1-1429) x CB519923 (1-743)

QY 1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243  
 DB 64 CTGTGTGTTTCAAAACCAACATCAAGGAATCAAGATAGTCTCTCTAATGAGAAATTT 123  
 QY 1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTriPro 1263  
 DB 124 GATTTTGGCAACATCCAAACATCACTTCTTGTGTTATGTCAGTAATTTCACTTGGCCC 183  
 QY 1264 IleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePhe 1283  
 DB 184 ATCAAAATTCAGATTGCCTCTCCAGCACACGAACTTCATGGACCTCGTACAGTTCTTC 243  
 QY 1284 ValThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTriLysIle 1303

Db 244 GTGACTTTCTTCAGTTGTGTTTCTCTCGCTGCTCTCGTGGCTGTCAGTGGTCTGGAAGATC 303  
 QY 1304 LysGlnSerCysTrpAlaSerArgArgGluGluGlnLeuLeuArgGlnMetGlnGlnMet 1323  
 Db 304 AGCAGAGCTGTGGGATCCAGCGGAGAGACAACTTCTTCGGAGATGCACAGATG 363  
 QY 1324 AlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProAsp 1343  
 Db 364 GCCAGCGCCCTTGTCTCTGTAACAGTTCCTTGGAAACAGATGAAGAACCTCTGAT 423  
 QY 1344 LeuIleGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGly 1363  
 Db 424 CTATTGGGGAGATATAAGACCGTCTCTTAACCCATTCCTGGAGCCCTGCTTGT 483  
 QY 1364 AsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyIlePro 1383  
 Db 484 AACAAAGCCGACGCTCTCTCTGTATTCTGTAGGCTCCCTCGAGGACTGGAGGAATCCCT 543  
 QY 1384 ProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMet 1403  
 Db 544 CCTCCTGGTCAGTCAGGTCCTGCTGTGGCCAGTGGCCCTGTGGACATTTCTCAGCAGATG 603  
 QY 1404 ProIleValTyrLysGluLysSerGlyValAlaValArgAsnArgLysGlnProAla 1423  
 Db 604 CCATAGTGTACAGGAGAGTCAGGAGCTGTGAAGAACCGGAGCAGCAGCGCTGCA 663  
 QY 1424 GlnProGlyThrCysIle 1429  
 Db 664 CAGCCTGGAACCTGCATT 681

RESULT 11  
 AY418589 3976 bp DNA linear GSS 17-DEC-2003  
 LOCUS Mus musculus ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,  
 DEFINITION genomic survey sequence.  
 AY418589  
 ACCESSION AY418589.1 GI:39774549  
 VERSION GSS.  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 3976)  
 REFERENCE Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous  
 gene trios.  
 TITLE Science 302 (5652), 1960-1963 (2003)  
 JOURNAL 14671302  
 PUBMED 2 (bases 1 to 3976)  
 REFERENCE Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
 Adams, M.D. and Cargill, M.  
 Direct Submission  
 TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
 JOURNAL Rockville, MD 20850, USA  
 COMMENT These sequences were made by sequencing genomic exons and ordering  
 them based on alignment.  
 FEATURES Location/Qualifiers  
 source 1..3976  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 <1..>3976  
 /gene="ATRN"  
 /locus\_tag="HCM6607"  
 gene  
 ORIGIN  
 Alignment Scores:

```

Pred. No.: 1,11e-155 Length: 3976
Score: 206.00 Matches: 576
Percent Similarity: 96.00% Conservative: 0
Best Local Similarity: 96.00% Mismatches: 12
Query Match: 14.42% Indels: 24
DB: 29 Gaps: 0

US-09-787-097-12 (1-1429) x AY418589 (1-3976)

Qy 842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861
Db 2210 ATGCAATCATCTCAAGATATGCTCAAGCTCACTGACTCCATGGGTGGTCTTCGGAAG 2269
Qy 862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881
Db 2270 ATCAATGCTTACTGCTGCTGGAGGATATGCTCCATTCACAAATAGTTTCTGCAG 2329
Qy 882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901
Db 2330 TGGATGCCATCTGAGCCAGTGATGCTGGCTTCTGTGGGATCTTGTGAGACCTAGTACT 2389
Qy 902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
Db 2390 CGGGGATTAAAGGCTGCAACCTGCAATCAACCTCTCAATGGCAGCGTCTGGAAGGCT 2449
Qy 922 AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp 941
Db 2450 GCACACACAGTGCACAGCAGTGCAGCACCATGTGCCCTCGGCACAGCGTGTGGCGA- 2508
Qy 942 -CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspS 961
Db 2509 GTGCACTAGCAG-CAGCTCGAGTGCATGTGGTGAGTAAATGATGATGATGAGTACT 2567
Qy 961 erAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerT 981
Db 2568 CCATATGCTTACTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2627
Qy 981 hr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000
Db 2628 G-CTGCCACCTGAAATATGCTCTGCTACTGTACTGCTGCTGCTGCTGCTGCTGCTGCT 2686
Qy 1001 GlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
Db 2687 GCTCTGTTGTGTACTGATCTTACCAATCTGGAAGAGGAAATGATGATGAGGCGCAGC 2746
Qy 1021 TyrLysGlyProValLysMetProSerGlnAlaProThr--GlyAsnPhe-TyrProGln 1039
Db 2747 TATAAGGACCTGTGAAGATGCGTCAACAGGCTC-TG-CAGGAAATGT-GTATCCACAG 2803
Qy 1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059
Db 2804 CCCCTTCTGAATCCAGCATGTGTCTAGAGGACAGCAGATACAACTGTGTCTTTCATTAC 2863
Qy 1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079
Db 2864 TGTCAGCTTTCAGTGTGCAACCGACACAGCAATGCATCAACAGATATCTGTGAGAAG 2923
Qy 1080 CysGluAsn-LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs 1099
Db 2924 TGTGAGGA-CCTGACACCGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGGTGA 2982
Qy 1099 pProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAs 1119
Db 2983 CCGGACTAATGGAGGCAATGTACGCAATGCAAGTGCATGCGCAGCATCTGTCGAA 3042
Qy 1119 nThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe 1139
Db 3043 CACCAACACCGGCAAGTCTTCTGTACCAACCAAGAGTGTCAAGGGGACAGTGCAGCT 3102
Qy 1139 uCysGluValGluAsnArgTyrGlnCysAsnProLeuArg-GlyThrCysTyrTyrThrL 1159
Db 3103 ATGTGAGGTAGAAATTCGATACAGGAAACCTCTCAA-AGGAACATGCTACTATATCCC 3161

```

```

Qy 1159 euLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrA 1179
Db 3162 TTCTCATTCAGTATCAGTTTACCTTTAGCTGTCCAGGAGACGACCGCTACTACACAG 3221
Qy 1179 laIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnA 1199
Db 3222 CCATCAACTTTGTGGCTACTCTCTGATGAACAAACAGGGATTGGACATGTTTCATCAATG 3281
Qy 1199 laSerLysAsnPheAsnLeuAsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrGl 1218
Db 3282 CTTCCAAAACCTTCAACCTCAACATCAGCTGGCCAC-CAGCTTCC-AGCGGAACCCA 3339
Qy 1218 nAla-GlyGluGluMet-ProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237
Db 3340 GAC-TGGAGAAAGAGT-GGCTGTGTGTTTCAAAAACCAACATCAGGAATACAAAGATAGC 3397
Qy 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257
Db 3398 TTCTCTAATGAGAAATTTGATTTTCCCAACCATCCAAACATCATTCTTCTTTGTTTATGTC 3457
Qy 1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277
Db 3458 AGTAATTTTCACTTGGCCCATCAAAATTCAGATTGCCCTTCTCCAGCACAGCAACTTCATG 3517
Qy 1278 AspLeuValGlnPhePheValThrPheSerCysPheLeuSerLeuLeuValAla 1297
Db 3518 GACCTGGTACAGTCTCTCGTGACTTTCTCAGTGTGTTTCTCTCGCTGCTTCTGGTGGCT 3577
Qy 1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeu 1317
Db 3578 GCAGTGGTCTGGAAGATCAAGCAGAGCTGTGCGGCATCCAGGCGGAGAGCAACTCTT 3637
Qy 1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337
Db 3638 CGGAGATGCAACAGATGCGCCGCCCTTTGCTTCTGTAAACCTTGCCTTTGGAACA 3697
Qy 1338 AspGluGluProAspLeuIleGlyGlySerIleLysThrValProLysProIleAla 1357
Db 3698 GATGAGAACCTCTGATCTCATTTGGGGAGTATTAAGACCTTCTTAAGCCCATTCGC 3757
Qy 1358 LeuGluProCysPheGlyAsnLysAlaValLeuSerValPheValArgLeuProArg 1377
Db 3758 CTGGAGCCCTGCTTGTGTAACAAAGCCGAGTCTCTCTGTATTCTGAGGCTCCCTCGA 3817
Qy 1378 GlyLeuGlyGlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal 1397
Db 3818 GGAATGGAGGAATCCCTCTCTGTCAGTCAGGCTCTGCTGTGGCCAGTGCCTGGTG 3877
Qy 1398 AspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArg 1417
Db 3878 GACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAGTCAAGAGCTGTGAAGAACCG 3937
Qy 1418 LysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db 3938 AAGCAGAGCGGCTGCACAGCCTGGAACCTGCATT 3973

```

```

RESULT 12
BE894518 601433126F1 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:3918271 5',
LOCUS mRNA sequence.
DEFINITION BE894518.1 GI:10356969
ACCESSION BE894518
VERSION BE894518.1
KEYWORDS EST.
SOURCE Homo sapiens (human)

```

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 633)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.

```





```

QY 1080 CysGluAenLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrglyAsp 1099
D 243 TGTGAGAACCTGACACAGGCAAGCACTGCAGACCTGCATATCTGGCTTCTACGGTGAT 302
QY 1100 ProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsn 1119
D 303 CCCACCAATGGAGGAATGTCAGCCATGCAAGTGCATGGCGCGCTCTGTGCAAC 362
QY 1120 ThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeu 1139
D 363 ACCAACCGGGCAAGTGTCTTCGCCACCAAGGGCGTCAAGGGGGAGCAGTGCACGTA 422
QY 1140 CysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyThrLeu 1159
D 423 TGTGAGGTAGAAATCGATACCAAGAAACCTCTCCAGAGNACATGTTATATACTCTT 482
QY 1160 LeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTyThrAla 1179
D 483 CTTATTACTATCAGTTCACCTTACCTTATAGTCTATCCAGGAAGATGATCGCTATTACAGCT 542
QY 1180 IleAsnPheValAlaThrProAspGluGlnAsnArgAspLeu-AspMetPheIleAsnAl 1199
D 543 ATCAATTTTGTGGCTACTCTCTGACCAACAAACAGGGATTC-GGACATGTTCAATCAATGC 601
QY 1199 aSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAl 1219
D 602 CTCCAAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 661
QY 1219 aGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSe 1239
D 662 TCGAGAAGAGATGCGCTGTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTCTC 721
QY 1239 rAsnGluLys 1242
D 722 TAATGAGAG 731

RESULT 14
LOCUS BM783821
DEFINITION K-EST0061830 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-H09 5',
mRNA linear EST 05-MAR-2002
ACCESSION BM783821
VERSION 1
KEYWORDS 1 (bases 1 to 677)
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 36 row: H column: 09
High quality sequence stop: 677.
FEATURES
Location/Qualifiers
1..677
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-36-H09"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"

```

```

/cell_line="SNU-484"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation
method."

ORIGIN
Alignment Scores: 4.44e-145 Length: 677
Pred. No.: 192.00 Matches: 224
Score: 99.12% Conservative: 0
Percent Similarity: 99.12% Mismatches: 1
Best Local Similarity: 13.44% Indels: 2
Query Match: 12 Gaps: 0
DB:

US-09-787-097-12 (1-1429) x BM783821 (1-677)
QY 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857
D 1 CAGCTGGCANTATGACGTATCTCAGAGCATCTCCNAGCTCACCCTTAACCCATGGGTC 60
QY 858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
D 61 GGCCTTCGGAAGATCAATGTGCTCTACTGTGTGGAGAGATATGTCCTCCCATTTACAAAT 120
QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
D 121 AGTTTACTAGTGGATGATGCGCTCTGAGCCAGTGTGATCTGGATTTCTGGAAATTTATCA 180
QY 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
D 181 GAACCCAGTACTCGGGACTGGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTGTC 240
QY 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
D 241 TGTGAAGGGCTGCAACACACAGTGTCTAAGCAGTGGCGGACACCATGTGCTTTGAGGACA 300
QY 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
D 301 GCATGTGAGATTCACACAGCGGAGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360
QY 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977
D 361 TGTGTGACTCCAATGCCTATGTGGCTCTCTCTTGGCCAGTGTATGGAATGGTAT 420
QY 978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
D 421 ACGATGAGCACCTGCCCTCCCTGAAATGTTCAGGCTACTGTACCTGTAGTCATGTTGTTG 480
QY 998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
D 481 GAGCAACACAGGCTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGAAATGCATA 540
QY 1018 GluGlySerTyrLysGlyProValLysMetProSerGln-AlaProThrGlyAsnPheTy 1037
D 541 GAGGTTTCTTATAAAGACCAAGTGAAGATGCTTCGNC-AGCCCCCTACAGAAATTTCTA 599
QY 1037 rProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPh 1057
D 600 TCCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTT 659
QY 1057 eIleHisCysProAla 1062

```

```

Db      660 CATTCTGTCAGCT 675
|||||
RESULT 15
BF672370
LOCUS   602150605F1 NIH_MGC_81 Homo sapiens cDNA clone IMAGE:4291680 5',
DEFINITION mRNA sequence.
ACCESSION BF672370
VERSION   BF672370.1 GI:11946265
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 917)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-research.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1136 row: j column: 01
High quality sequence stop: 563.
FEATURES
source
1..917
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4291680"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 81"
/note="Organ: muscle (skeletal); Vector: pDNR-LIB
(Clontech); Site 1: SfiI (ggccgctggcc); Site 2: SfiI
(ggcatatggcc); 5' and 3' adaptors were used in cloning
as follows: 5' adaptor sequence: 5'-CACGGCCATATGACC-3'
and 3' adaptor sequence:
5'-ATTCATGAGCGCGCGCCGACATG-dT(30)EN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.55 kb (range 1.0-4.0 kb). 15/15 colonies contained
inserts by PCR. This library was enriched for full-length
clones and was constructed by Clontech Laboratories (Palo
Alto, CA)."
```

## ORIGIN

```

Alignment Scores:
Pred. No.:      2,82e-139      Length:      917
Score:          185.00         Matches:    185
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match:     12.95%         Indels:    0
DB:              10            Gaps:      0
```

US-09-787-097-12 (1-1429) x BF672370 (1-917)

```

QY      945 GlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyr 964
Db      1   GGAGGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
QY      965 ValAlaSerPheProGlyGlnCysMetGluTrpTyrThrMetSerThrCysPro 984
Db      61   GTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY      985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnProGlyCysGlyTyr 1004
Db      121   GAAATTTGTTACGGCTACTGTACCTGTACCTGTACCTGTACCTGTACCTGTACCTGTACCTGTAC 180
```

```

QY      1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024
Db      181   TGTACTGATCCCGCAATACTGGCAAGGGAATGATAGAGGGTTCCTATAAGGACCA 240
QY      1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044
Db      241   GTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGTCTCAATTC 300
QY      1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGln 1064
Db      301   AGCATGTGTCTAGAGGACAGCAGATCAACTGTCTTTTCAATTCACCTGTGACGCTTCCCAA 360
QY      1065 CysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThr 1084
Db      361   TGCACGGCCACAGTAAATGATCAATCAAGCATCTGTGAGAAGTGTGAGAAGTGTGAGAAGTGT 420
QY      1085 ThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGly 1104
Db      421   ACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGTATCCCAATGAGGG 480
QY      1105 LysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLys 1124
Db      481   AATGTGAGCATGCAAGTCAATGGCAGCCGCTCTCTGTGCAACACCAACACGGGCAAG 540
QY      1125 CysPheCysThrThr 1129
Db      541   TGCTTCTGCACCACC 555

RESULT 16
BF783788
LOCUS   K-EST0061794 SSSNU484 Homo sapiens cDNA clone SSSNU484-36-E09 5',
DEFINITION mRNA sequence.
ACCESSION BF783788
VERSION   BF783788.1 GI:19132020
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 683)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE   21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoseun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 36 row: E column: 09
High quality sequence stop: 683.
FEATURES
source
1..683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SSSNU484-36-E09"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F"
/clone_lib="SSSNU484"

```

vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F<sup>+</sup> by electroporation method."

## ORIGIN

Alignment Scores:  
Pred. No.: 4,74e-128 Length: 683  
Score: 171.00 Matches: 225  
Percent Similarity: 98.25% Conservative: 0  
Best Local Similarity: 98.25% Mismatches: 2  
Query Match: 11.97% Indels: 4  
DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM783788 (1-683)

```
QY 838 GlnLeuArglleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTIPVal 857
DB 1 CAGCTGCGGAATTAATGAGTATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGGTC 60

QY 858 GlyLeuArgLysIleAsnValSerTyrTyrCysTyrGluAspMetSerProPheThrAsn 877
DB 61 GGCCTTCGGAAGATCAATGTGCTCTACTGCTGGGAGATATGTCGCCATTTACAAAT 120

QY 878 SerLeuLeuGlnTyrMetProSerGluProSerAspAlaGlyPheCysGlyIleuSer 897
DB 121 AGTTACTACAGTGGATCGGCTCTGAGCCAGTGTGCTGATCTGCGAATTTATCA 180

QY 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
DB 181 GAACCCAGTACTCGGGAGTGAAGCTGCAACCTGCAACCCACATCAATGGTAGTGC 240

QY 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
DB 241 TGTGAAGGCTGCAACACAGTCTAAGCAGTCCGACACCATGTCCTTGAGGACA 300

QY 938 AlaCysGlyAspCysThrSerGlySerGluCysMetTyrCysSerAsnMetLysGln 957
DB 301 CGATGTGAGATGTCACCAAGCGGAGCTCTGATGTCATGTTGGTGGAGCAATGAGCAG 360

QY 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrTyr 977
DB 361 TGTGTGGACTCAATGCTATGTGGCTCTTCCCTTTGGCCAGTGTATGGAATGGTAT 420

QY 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
DB 421 ACGATGAGCAGCTGCCCCCTGAAATTTGTTGAGGCTACTGTACCTGTAGTCAITGCTTG 480

QY 998 GluGlnProGlyCysGlyTyrCysThrAspProSerAsn--ThrGlyLysGlyLysCysI 1017
DB 481 GAGCAACACAGGCTGTGGCTGTGTACTGATCCAG-NC-ATACTGGCAAGGGAATGCA 538

QY 1017 leGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheT 1037
DB 539 TAGAGGGTTTCCTATTAAGGACCCAGTGAAGATGCTTCGCAAGCCCTCAGAGAAATTTCT 598

QY 1037 yProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTyrSerP 1057
DB 599 ATCCAGGCCCTGTCTCAATTCACGATGTGTCTAGAGGACAGAGATACAACTGGTCTT 658

QY 1057 heileHisCysProAlaCysGln 1064
DB 659 TCATTCAGTGTCCAGCTTGCCAA 681
```

## RESULT 17

BM772968 98.25% 486 bp mRNA linear EST 04-MAR-2002  
LOCUS K-EST0057334 S5SNU484s1 Homo sapiens cDNA clone S5SNU484s1-3-A12  
DEFINITION 5', mRNA sequence.  
ACCESSION BM772968

BM772968.1 GI:19102583

## KEYWORDS

EST. Species (human)

## SOURCE

Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 486)

## AUTHORS

Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,

Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and

Kim,Y.S.

## TITLE

21C Frontier Korean EST Project 2001

## JOURNAL

Unpublished (2002)

## COMMENT

Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience &amp; Biotechnology

52 Eeum-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 3 row: A column: 12

High quality sequence stop: 486.

## FEATURES

source

1..486

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S5SNU484s1-3-A12"

/sex="M"

/tissue\_type="Stomach"

/cell\_type="Epithelial"

/cell\_line="SNU-484"

/lab\_host="Top10F"

/clone\_lib="S5SNU484s1"

/notes="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;

Site 2: NotI; The poly (A)+ RNA was decapped with tobacco

acid pyrophosphatase (TAP) and ligated with DNA-RNA linker

including EcoRI site by treatment of T4 RNA ligase. The

first strand cDNA was synthesized from oligo dT-selected

mRNA by priming with dT-tailed vector. The dT-tailed

vector was adjusted to have about 60nt. The cDNA vector

was circularized with E. coli DNA ligase after digestion

of EcoRI which site is also included in vector. An RNA

strand converted to a DNA strand by Okayama-Berg method.

The obtained cDNA vectors were used for transformation of

competent cells E. coli Top10F<sup>+</sup> by electroporation method.

After analyzing and sequencing about 2,000 ~ 3,000

colonies in original cDNA library, the abundant cDNAs were

selected and amplified by PCR reaction using vector region

primer including T7 promoter as 5' primer and N(dT)14 as

3' primer. The PCR products were used as template for

synthesis of biotinylated single stranded RNA by in vitro

transcription reaction. The synthesized RNA probes were

hybridized with antisense single stranded cDNAs prepared

from original library and incubated with avidin-gel.

After removing DNA-RNA hybrids by centrifuge, the

subtracted cDNA libraries were constructed by

transformation of the remaining DNA into competent cells E.

coli Top10F<sup>+</sup> with electroporation method."

## ORIGIN

## Alignment Scores:

Pred. No.: 4,35e-120 Length: 486  
Score: 161.00 Matches: 161  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.27% Indels: 0  
DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM772968 (1-486)

QY 838 GlnLeuArglleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTIPVal 857

DB 1 CAGCTGCGGAATTAATGAGTATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGGTC 60

QY 858 GlyLeuArgGlyIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877  
 Db 61 GGCCCTTCGGAAGATCAATGTCTCTACTGGTGGGAAGATATGTCCCAATTTACAAT 120  
 QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897  
 Db 121 AGTTTACTACAGTGGATGGCGCTGAGCCAGTGGATCTGGATCTGGTGGATTTATCA 180  
 QY 898 GluProSerThrArgGlyLeuIysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917  
 Db 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAAGCTGCAATCAACCCACTCAATGGTAGTGT 240  
 QY 918 CysGluArgProAlaAsnHisSerAlaIysGlnCysAspThrProCysAlaLeuArgThr 937  
 Db 241 TGTGAAGGCTTGCAACACAGCTGTAAGCAGTGGCCGACACCATGTGCTTGAGGACA 300  
 QY 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957  
 Db 301 GCATGTGGAGATTGCACACGCGCAGCTCTGAGTGCATGCTGGTGAGCAACATGAAGCAG 360  
 QY 958 CysValAspSerAsnAlaTrpValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977  
 Db 361 TGTGTGGATCTCAATGCCCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTAT 420  
 QY 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997  
 Db 421 ACGATGAGCACCTGCCCTCTGAAATTTGTTCCAGCTACTGTACCTGTAGTCAATGCTTG 480  
 QY 998 Glu 998  
 Db 481 GAG 483

RESULT 18  
 BX431593  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 962)  
 Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 10212.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 c91-bin/cluster.cgi?seq=CSOBAG0172B02\_CS01546\_1&cluster=10212.f.  
 Contact : Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/InvitrogenCorporation1600  
 Faraday Avenue Genoscope sequence ID : CSOBAG0172B02\_CS01546\_1.  
 Location/Qualifiers  
 1. 962  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSODF012YL19"  
 /tissue type="FETAL BRAIN"  
 /dev stage="fetal"  
 /clone.lib="Homo sapiens FETAL BRAIN"  
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo (GT) primer. Five prime end  
 enriched, double-strand cDNA was digested with Not I and

FEATURES  
 source

cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
 vector. Library was not normalized."

ORIGIN  
 Alignment Scores:  
 Pred. No.: 6,19e-115 Length: 962  
 Scores: 155.00 Matches: 204  
 Percent Similarity: 99.03% Conservatives: 0  
 Best Local Similarity: 99.03% Mismatches: 1  
 Query Match: 10.85% Indels: 2  
 DB: 13 Gaps: 0  
 US-09-787-097-12 (1-1429) x BX431593 (1-962)  
 QY 1054 AsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsn 1073  
 Db 53 AACTGGTCTTTCAATTCACCTGCCAGCTTCCCAATGCAACGGCCACAGTAATCATCAAT 112  
 QY 1074 GlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIle 1093  
 Db 113 CAGGACATCTGTGAGAGTGTGAGAACCTTGACCAAGGACAGCACTGCCGAGACCTGCATA 172  
 QY 1094 SerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGly 1113  
 Db 173 TCTGGCTTCTACGGTGATCCCAATGGAGGGAATGTCCAGCATGCAAGTCAATGGG 232  
 QY 1114 HisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLys 1133  
 Db 233 CACGGTCTCTGTGCAACACCAACACGGCAAGTGTCTTGACACCAAGGGCGTCAAG 292  
 QY 1134 GlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGly 1153  
 Db 293 GGGGACAGTGGCCAGCTATGTGAGGTAGAAATTCGATACCAAGGAAACCTCTCGAGGA 352  
 QY 1154 ThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAsp 1173  
 Db 353 ACATGTATTATTAATCTCTTCTATTGACTACTAGTTCACCTTTAGTCTATCCCAGGAAGAT 412  
 QY 1174 AspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeu 1193  
 Db 413 GATCGCTATTACACAGCTATCAATTTTGTGGTACTCTCTGACGACACAAACAGGATTTG 472  
 QY 1194 AspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrp-AlaAlaSerPh 1213  
 Db 473 GACATGTTTCATCAATGCCCTCCAGGATTTCAACCTCAACATCACCNT-GGCTGCCAGTTT 531  
 QY 1213 eSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGln 1233  
 Db 532 CTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAAGA 591  
 QY 1233 uTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPh 1253  
 Db 592 GTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGACCAACCAACCAATATCATCTT 651  
 QY 1253 ePheValTyrValSer 1258  
 Db 652 CTTTGTATTATGTCAGT 667

RESULT 19  
 BX431593  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 744)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 BQ183626 744 bp mRNA linear EST 15-JUL-2003  
 UI-H-EUO-azn-a-22-0-UI-s1 NCI-CGAP\_Carl Homo sapiens cDNA clone  
 IMAGE:5850909 3', mRNA sequence.  
 BQ183626  
 BQ183626.1 GI:20359183  
 EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 744)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Jose Mercuende  
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
The following repetitive elements were found in this cDNA  
sequence: 16-237, >11P22#LINE/L1 (matched complement)  
Seq primer: M13 FORWARD  
POLYA=Yes.

Location/Qualifiers  
1..744  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5850909"  
/tissue\_type="Osteoarthritic Cartilage"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NCI CGAP Carl"  
/note="Organ: Knees; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Carl is a cDNA library containing the following tissue(s): Osteoarthritic Cartilage. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT.  
TAG\_TISSUE=osteoarthritic cartilage  
TAG\_LIB=UI-H-EUO  
TAG\_SEQ=TGATCAGCT"

ORIGIN

Alignment Scores:  
Pred. No.: 3,12e-114 Length: 744  
Score: 154.00 Matches: 154  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.78% Indels: 0  
DB: 13 Gaps: 0

US-09-787-097-12 (1-1429) x BQ183626 (1-744)

QY 1114 HisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrIysGlyValIys 1133  
Db 697 CACGCGTCTGTGCAACCAACACGCGGCAAGTGTCTTGACCACCAAGGGCGTCAAG 638  
QY 1134 GlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGly 1153  
Db 637 GGAGACGAGTGCAGCTATGTAGGTAGGATGAGAAATCGATACAGGAACCTCTCAGAGGA 578  
QY 1154 ThrCysTyrThrLeuLeuLeuAspTyrGlnPheThrPheSerIeuSerGlnGluAsp 1173  
Db 577 ACATGTTATTACTCTCTTTTGTACTATCAGTTCACCTTTAGTCTATCCAGGAAGAT 518  
QY 1174 AspArgTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeu 1193  
Db 517 GATCGCTATTACACGCTATCAATTTTGTGGCTACTCTCGACGACCAAAACAGGGATTG 458  
QY 1194 AspMetPheIleAsnAlaSerIysAsnPheAsnLeuAsnIleThrTrpAlaIaSerPhe 1213  
Db 457 GACATGTTTCATCATGCTCCAGAGATTTCAACCTCAACATCACCTGGGCTGCCAGTTTC 398

QY 1214 SerAlaGlyThrGlnAlaGlyGlnGluMetProValValSerIysThrAsnIleLysGlu 1233  
Db 397 TCAGCTGGAAACCCAGGCTGGAGAGATGCTGTGTTTCAAAAACCAACATTAGGAG 338  
QY 1234 TyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhe 1253  
Db 337 TACAAAGATAGTTCCTCTAATGAGAAGTTGATTTTCGCAACCAACCAATATCACTTTC 278  
QY 1254 PheValTyrValSerAsnPheThrTrpProIleLysIleGln 1267  
Db 277 TTGTTTATGTCAGTAATTCACCTGCCCATCAAAATTCAG 236

RESULT 20  
BF853962  
LOCUS BF853962 544 bp mRNA linear EST 16-JAN-2001  
DEFINITION MR2-EN0093-261200-004-g08-EN0093 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF853962  
VERSION BF853962.1 GI:12241706  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 544)  
AUTHORS Dias Neto, E., da Silva, W. Jr., Garcia Correa, R., Verjovski-Almeida, S., Brines, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PubMed 10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-261200-004-g08&t3=2000-12-26&t4=1>)  
Seq primer: puc 18 forward  
High quality sequence stop: 544.  
Location/Qualifiers  
1..544  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="EN0093"  
/note="Organ: lung normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Alignment Scores:  
Pred. No.: 4.02e-111 Length: 544  
Score: 150.00 Matches: 150  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.50% Indels: 0  
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BF853962 (1-544)

```

QY 1113 GlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyVal 1132
Db 64 GGCACGCGTCTCTGTGCAACACACAGCGGAGTGTCTTGCCACCAAGGCGTC 123
QY 1133 LysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg 1152
Db 124 AAGGGGACAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAACCTCTCAGA 183
QY 1153 GlyThrCysTyrTyrThrLeuLeuLeuAspTyrGlnPheThrPheSerLeuSerGlnGlu 1172
Db 184 GGAACATGTTACTATCTCTCTTATGACTATACAGTTCACCTTGTAGTCTATCCAGGAA 243
QY 1173 AspAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAsp 1192
Db 244 GATGATCGCTATTACACAGCTATCAATTTTGGCTACTCTCTGACGAACAAACAGGGAT 303
QY 1193 LeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaIleAsr 1212
Db 304 TTGGACATGTTTCATCAATCCTCCCAAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGT 363
QY 1213 PheSerAlaGlyThrGlnAlaGlyGluMetProValValSerLysThrAsnIleLys 1232
Db 364 TTCTCAGCTGGACCCAGCTCGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAAG 423
QY 1233 GluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThr 1252
Db 424 GAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTGGCAACCAACCAATATCACT 483
QY 1253 PhePheValTyrValSerAsnPheThrTrp 1262
Db 484 TTTCTTTGTTATGTCAGTAATTTACCTGG 513

```

## RESULT 21

CB242644/c

LOCUS

DEFINITION UI-CF-FNO-afs-i-02-0-UI s1 UI-CF-FNO Homo sapiens cDNA clone

UI-CF-FNO-afs-i-02-0-UI 3', mRNA sequence.

CB242644

CB242644.1 GI:28364288

EST.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 746)

Bonaldo, M.F., Lennon, G. and Soares, M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

8889548

Contact: McCray, PB

McCray Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com) or from Open Biosystems

(www.openbiosystems.com).

The following repetitive elements were found in this cDNA

sequence: 16-237, &gt;L1PA2#LINE/L1 (matched complement)

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

1. .746

source

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="UI-CF-FNO-afs-i-02-0-UI"

/tissue\_type="Human Lung Epithelial cells"

/lab\_hosts="DH10B (Life Technologies) (T1 phage resistant)"

/clone\_lib="UI-CF-FNO"

/notes="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-CF-FNO is a subtracted cDNA library derived from two

normalized human lung epithelial cell libraries (EN1 and

DUI) The library was subtracted according to according to

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. For additional information, contact:

bento-soares@uiowa.edu

TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS

6hr to LPS 24h

TAG\_LIB=UI-CF-FNO

TAG\_SEQ=CTGCTCAGGT"

## ORIGIN

Alignment Scores:

Pred. No.: 1,67e-104 Length: 746

Score: 142.00 Matches: 169

Percent Similarity: 99.41% Conservative: 0

Best Local Similarity: 99.41% Mismatches: 1

Query Match: 9.94% Indels: 1

DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CB242644 (1-746)

```

QY 1098 GlyAspProThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeu 1117
Db 744 GGTGATCCCAATGAGGAAATGTCAGCCATGCAAGTGCATGCGGACGCGTCTCTG 685
QY 1118 CysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCys 1137
Db 684 TGCACACCAACACGGGCAAGTG-TTCTGCACCAACAGGGCGTCAAGGGGACGAGTGC 626
QY 1138 GlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyr 1157
Db 625 CAGCTATGTGAGGTGAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTAT 566
QY 1158 ThrLeuLeuLeuAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTyr 1177
Db 565 ACTCTTCTTATTGACTATCAGTTTACCTTTAGTCTATCCAGAGAGATGATCGCTATTATAC 506
QY 1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197
Db 505 ACAGCTATCAATTTTGTGGCTACTCTCTGACGAACAAACAGGGATTTGGACATGTTATC 446
QY 1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThr 1217
Db 445 AATGCCTCCAGAAATTTCAACCTCAACATCACCTGGCTGCCAGTTTCTCAGCTGGAACC 386
QY 1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237
Db 385 CAGCTCGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGT 326
QY 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257
Db 325 TTCTCTTAATGAGAAGTTTGATTTTCGCAACCAACCAATATCATTCTTCTTTTATGTC 266
QY 1258 SerAsnPheThrTrpProIleLysIleGln 1267
Db 265 AGTAATTTCACTGGCCCATCAAAATTTCAG 236

```

## RESULT 22

BI850289

LOCUS

DEFINITION

imageqc.11.2000/sly246bdr81.v1 NIH\_MGC\_56 Homo sapiens cDNA clone

IMAGE:4284387 5', mRNA sequence.

ACCESSION

BI850289

```

VERSION BI850289.1 GI:16003776
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 591)
AUTHORS Kale, P.I., Harsch, T.J., Folta, P.A., Nelson, D.O., Sanders, C.G. and
Prange, C.K.
TITLE The I.M.A.G.E. Consortium quality control effort: clone
resequencing for verification
JOURNAL Unpublished (2001)
COMMENT Other_ESTS: BF700035
Contact: Prange CK
The I.M.A.G.E. Consortium
Lawrence Livermore National Laboratory
Livermore, CA, USA
Email: help@image.llnl.gov
This read has been verified (found to hit its original self in the
correct orientation), as part of the I.M.A.G.E. Consortium quality
control effort. High quality sequence is defined as having 100 or
more base pairs with a phred quality value of 20 or greater, where
a sliding window of 4 base pairs marks the beginning and end of the
sequence. For information on obtaining this clone, please contact
info@image.llnl.gov.
Plate: LLCM117 row: j column: 4
Seq primer: -21m13
High quality sequence stop: 591.
Location/Qualifiers
1..591
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4284387"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (TI phage-resistant)"
/clone_lib="NIH MGC 56"
/notes="Organ: brain; Vector: pBMR-LIB (Clontech); Site 1:
SFII (ggccctcgcc); Site 2: SFII (ggccattatggcc);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-ATTGAGCGCGGCGGCACATG-dt(30)BN-3'
(where B = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
```

## FEATURES

source

## ORIGIN

```

Alignment Scores:
Pred. No.: 8,62e-104 Length: 591
Score: 141.00 Matches: 195
Percent Similarity: 99.49% Conservatives: 0
Best Local Similarity: 99.49% Mismatches: 1
Query Match: 9.87% Indels: 1
DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BI850289 (1-591)

QY 880 LeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluPro 899
Db 3 CTACAGTGGATGCGCTGAGCGCCAGTGTGGATTCTGTGGAAATTTATCAGAACCC 62
QY 900 SerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGlu 919
Db 63 AGTACTCGGGAGCTGAGGCTGCAACTGCATCAACCCACTCAATGTTAGTCTGTGAA 122
QY 920 ArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCys 939
Db 123 AGGCTCTGCAACACACAGTGTACAGCAGTGCAGCAGACACCATGTGC-CTGAGGACAGATGT 181
QY 940 GlyAspCysThrSerGlySerGluCysMetTrpCysSerAsnMetLysGlnCysVal 959
```

```

Db 182 GGAGATTGCACACGCGCAGCTCTGAGTCATGTGGTGACCAACATGAAGCAGTGTG 241
QY 960 AspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMet 979
Db 242 GACTCCAAATGCCATATGTGGCCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATG 301
QY 980 SerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGln 999
Db 302 AGCACCTGCCCCCTGAAAATTGTTGAGCTACTGTACCTGTAGTACCTGTGTGGAGCAA 361
QY 1000 ProGlyCysGlyTrpCysThrAspProSerAsnThrGlyGlyGlyCysIleGluGly 1019
Db 362 CCAGGCTGTGGTGTGTACTGATCCAGCAATACTGCGCAAAAGGGAATGCATAGAGGT 421
QY 1020 SerTyrIysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGln 1039
Db 422 TCTATAAAGGACAGTGAAGATGCTTCGAGGCCCTTACAGAAATTTCTATCCACAG 481
QY 1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059
Db 482 CCCTGCTCAATTCAGCATGTCTTAGAGGACAGCAGATACAACTGGTCTTTCAATTCAC 541
QY 1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSer 1075
Db 542 TGTCCAGCTTGCCATGCACAGGCCACAGTAATGCATCAATCAGAGC 589
```

## RESULT 23

CA306924/c

LOCUS

DEFINITION

UI-H-FTI1-bhu-e-21-0-UI.s1 NCI\_CGAP\_FTL Homo sapiens cDNA clone

CA306924

ACCESSION

VERSION

CA306924.1

GI:24469978

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 643)

AUTHORS

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapsb-x@mail.nih.gov

Tissue Procurement: Dr. Gary W. Hunninghake, U of I

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bento-soares@uiowa.edu

The following repetitive elements were found in this cDNA

sequence: 16-237, &gt;L1PA2#LINE/L1 (matched complement)

Seq primer: M13 FORWARD

POLYA-Yes.

Location/Qualifiers

1..643

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="UI-H-FTI1-bhu-e-21-0-UI"

/tissue\_type="Aveolar Macrophage"

/dev\_stage="adult"

/lab\_host="DH10B (Life Technologies)"

/clone\_lib="NCI\_CGAP\_FTL"

/note="Organ: Lung; Vector: p7T73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

NCI CGAP FTL is a normalized cDNA library constructed from

a pool of 81 RNA samples from Alveolar Macrophages

challenged with different treatments. The library was

normalized according to Bonaldo, Lennon and Soares, Genome

Research, 6:791-806, 1996. First strand cDNA synthesis was

primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pVT3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GGCATGCGG. The tissue was provided by Dr. Gary W. Hunninghake of the University of Iowa.

TAG\_LIB=UI-H-FTI  
TAG\_SEQ=GGCCATGCGG

## ORIGIN

## Alignment Scores:

Pred. No.: 1,06e-99 Length: 643  
Score: 136.00 Matches: 136  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.52% Indels: 0  
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CA306924 (1-643)

QY 1132 ValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrcGlnGlyAsnProLeu 1151  
DB 643 GTCAAGGCGNACGAGTGCCAGCTATGTGAGGTAGAAATCGATACCAAGGAACCTCTC 584  
QY 1152 ArgGlyThrCysTyrrTyrrThrLeuLeuLeuAspTyrrGlnPheThrPheSerLeuSerGln 1171  
DB 583 AGAGGAACATGTTATTATATCTCTTATTGACTATCATCAGTTACCTTTAGTCTATCCCGAG 524  
QY 1172 GluAspAspArgTyrrTyrrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArg 1191  
DB 523 GAAGATGATCGCTATTACACAGCTATCAATTTTGGCTACTCTCGACGACAAACAGG 464  
QY 1192 AspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAla 1211  
DB 463 GATTTGGACATGTTTCATCAATGCCCTCCAAAGATTTCACCTCAACATCACCTGGGCTGCC 404  
QY 1212 SerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIle 1231  
DB 403 AGTTTCTCAGCTGGAAACCCAGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATT 344  
QY 1232 LysGluTyrrLysAspSerPheSerAsnGlnLysPheAspPheArgAsnHisProAsnIle 1251  
DB 343 AAGGAGTACAAAGATGTTCTCTAATGAGAAGTTTGATTTCGCAACCAACCCCAATATC 284  
QY 1252 ThrPhePheValTyrrValSerAsnPheThrTrpProIleLysIleGln 1267  
DB 283 ACTTTCTTGTTTGTGTCAGTAATTTTCACTGGCCCAATCAAAATTCAG 236

## RESULT 24

## BE880905

## LOCUS

DEFINITION 601490448P1 NIH\_MGC\_69 Homo sapiens cDNA clone IMAGE:3892859 5',  
mRNA sequence.

VERSION BE880905

KEYWORDS BE880905.1 GI:10329681

SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 991)

NIH-MGC <http://mgs.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)

Tissue Procurement: DCTD/DP/Gapdar

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.W.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be  
found through the I.W.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: LAM9679 row: p column: 12

High quality sequence stop: 607.

## FEATURES

## source

1..991  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3892859"  
/tissue\_type="large cell carcinoma, undifferentiated"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_69"  
/notes="Organ: lung; Vector: pCMV-SPORT6; Site\_1: NotI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.1 kb. Library constructed by Life  
Technologies."

## ORIGIN

## Alignment Scores:

Pred. No.: 1,61e-99 Length: 991  
Score: 136.00 Matches: 176  
Percent Similarity: 98.32% Conservative: 0  
Best Local Similarity: 98.32% Mismatches: 1  
Query Match: 9.52% Indels: 3  
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BE880905 (1-991)

QY 513 AlaLeuTyrrValHisGlyGlyTyrrLysAlaPheSerAlaAsnLysTyrrArgLeuAlaASP 532  
DB 39 GCCTATACGTTTCATGTTGGCTACAAAGCTTTTCAGTGCCTCAATAAGTACCGCTTCGAGAT 98  
QY 533 AspLeuTyrrArgTyrrAspValAspThrGlnMetTyrrThrIleLeuLysAspSerArgPhe 552  
DB 99 GATCTCTACCGATGATGTGGATACCCAGATGTGCACCATCTTCTTAAAGACACGCCATT 158  
QY 553 PheArgTyrrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsn 572  
DB 159 TTCCGTACTTGCACACAGCTGTGATGATGAGTGGACCATCTGCTGTGTTGAGGAAC 218  
QY 573 ThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAla 592  
DB 219 ACACACAATGACACATCTATGAGCCATGGCGCAAAATGCTTCTTTCAGATTTTCATGGCC 278  
QY 593 TyrrAspIleAlaCysAspArgTyrrSerValLeuProAspProAspLeuHisAspVal 612  
DB 279 TATGACATTGCCGTGACCGCTGTCAGTCTTCCAGACCTGATCTCCACCATGATGTC 338  
QY 613 AsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyrrValPheGlyGlyPhe 632  
DB 339 AACAGATTTGGCCATTTCAGCAGCTTACACAACAGCACCATGATGTGTGTTGCGTGTTC 398  
QY 633 AsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGluGln--CysAspAlaHisA 652  
DB 399 AATAGTCTCTCTCTCAGCGACATCTCTGGTATTCACCTCGGAACAGTGTGCAT--GGCATT 457  
QY 652 rgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlyS 672  
DB 458 GGAGTGAACCCGCTTTTGTAGCAGCAGGACCTGGTATTTCGGTGTGTGTGGAACACAGG 517  
QY 672 erSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGlnGlnLysLeu 689  
DB 518 CGTCTCAGTGTATCTCTGTGGCGCTGGCAACTGATGAACAAGAAAGATTGA 570

## RESULT 25

## AY418588

## LOCUS

Pan troglodytes ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,

Genomic survey sequence.

ACCESSION AY418588

AY418588 3671 bp DNA linear GSS 17-DEC-2003



**VERSION** AY418588.1 GI:39774548  
**SOURCE** GSS.  
**ORGANISM** Pan troglodytes (chimpanzee)  
Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.

---

**REFERENCE**  
**AUTHORS** Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,  
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,  
Adams,M.D. and Cargill,M.  
Infering nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
Science 302 (5652), 1960-1963 (2003)

---

**JOURNAL** PUBLISHED 14671302  
**REFERENCE** 2 (bases 1 to 3671)  
**AUTHORS** Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,  
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,  
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,  
Adams,M.D. and Cargill,M.  
Direct Submission  
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
These sequences were made by sequencing genomic exons and ordering  
them based on alignment.

---

**TITLE**

---

**FEATURES**  
source Location/Qualifiers  
1..3671  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
<1..>3671  
/gene="ATRN"  
 locus\_tag="HCM6607"

---

**gene**

---

**ORIGIN**

---

Alignment Scores:			
Pred. No.:	3,74e-98	Length:	3671
Score:	135.00	Matches:	229
Percent Similarity:	98.28%	Conservative:	0
Best Local Similarity:	98.28%	Mismatchches:	2
Query Match:	9.45%	Indels:	4
DB:	29	Gaps:	0

---

US-09-787-097-12 (1-1429) x AY418588 (1-3671)

---

Oy	1158	ThrLeuLeuleAspTyrGlnPheThrPhSerIeuSerGlngluAaspArgTyTYr	1177
Dd	2978	ACTCTTCATTAGTAGTTACGTCCACTTGCTATGCCAGGAAGATGCCTATTAC	3037
Oy	1178	ThrAlalleAsnPheValAlaThPrOspGluglNasnArgasPlieuAspMetPheIlle	1197
Dd	3038	ACAGCTATTAAATTTGGTGCTACTCCTCGACAACAAAACAGGGATTGGACATGTTCA TC	3097
Oy	1198	AsnAlaserLyssAsnPheAsnLsuAnlleThrrTrpalalaserPheSerlaGIlythr	1217
Dd	3098	AATGCCTCCAAGAATTTCAA CTCAACATCACCTGGGCTGCCAGTTCTCAGCTGGAAC C	3157
Oy	1218	GlnAlaglyGIUGluMeTProvalValserLystrAsnlleyysGIutyrlYlsAspSer	1237
Dd	3158	CAGCTCGAAGAAGATGCCTGTTGTTTTCA AAAACCACATT AAGAGTACAAAGATAG T	3217
Oy	1238	PheSerAsngLUylsyPheAsppheaRgasNHisProsnlleThrphepheValtyrVal	1257
Dd	3218	TTCTCTAATGAAGAAGTTTGATTTTCG CAACACCAATA ATCATCTTCTTTGTTATGTC	3277
Oy	1258	SerAsnPheThrTpTroillelysllledlnllealaPe--SerGlnHisSerAsnpHeM	1277
Dd	3278	AGTAATTTCACTGCCCCCATCAAAATTCAGATTGCC TTCTCTCNN--AC-ACGAATTT TA	3335
Oy	1277	e tAspLeuVaIGlnPhePheValThrphePheSerCy sPheLeuseri leuLeuVala	1297
Dd	3336	TGGACCTGGTACAGTTCTTCGTGACTTTCTTTCAGTTGTTTCCTCTCTTTGCTCTG GTG	3395

## Alignment Scores:

Pred. No.: 3,87e-98 Length: 560  
 Score: 134.00 Matches: 134  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 9.38% Indels: 0  
 DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BF853965 (1-560)

QY 1113 GlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyVal 1132  
 DB 62 GGCACGGCTCTCTGTGCAACACACACGGCAAGTGTCTTGACACACCAAGGGCTC 121  
 QY 1133 LysGlyAspGluCysGlnLeuCysGluValGluAsnArgTy-GlnGlyAsnProLeuArg 1152  
 DB 122 AAGGGGGACGAGTCCAGCTATGTGAGGTAGAAAATCGATACCAAGAAACCTCTCAGA 181  
 QY 1153 GlyThrCysTyThrThrLeuLeuLeuAspTyThrPheThrPheSerLeuSerGlnGlu 1172  
 DB 182 GGNACATGTTATTAATCTCTCTTATGACTACCACTTACCTTTAGTCTATCCAGGAA 241  
 QY 1173 AspAspArgTyThrThrAlaLeuAsnPheValAlaThrProAspGluGlnAsnArgAsp 1192  
 DB 242 GATGATCGCTATTACACAGCTATCAATTTGTGGCTACTCTCTGACCAACAAACAGGGAT 301  
 QY 1193 LeuAspMetPheLeuAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSer 1212  
 DB 302 TTGGACATGTTTATCAATGCTCCAGAAATTTCAACCTCAACATCACTGGGCTCCAGT 361  
 QY 1213 PheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLys 1232  
 DB 362 TTCTAGCTGGAAACCCAGCTGAGAGAGATGCTGTGTTTCAAAACCAACATTAG 421  
 QY 1233 GlyTyThrLysAspSerPheSerAsnGluLysPheAspPheArg 1246  
 DB 422 GAGTCAAAAGATAGTTTCTTAATGAGAGATTGATTTCGC 463

## RESULT 27

LOCUS CB521623 782 bp mRNA linear EST 09-JUL-2003  
 DEFINITION UI-R-GH0-cen-m-06-0-UI.r1 NIH\_BMAP\_GH0 Mus musculus cDNA clone  
 IMAGE:6841999 5', mRNA sequence.  
 ACCESSION CB521623  
 VERSION CB521623.1 GI:29354978  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NIH-MGC <http://mgi.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: Dr. Jim Lin, University of Iowa  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
<http://genome.uiowa.edu/distribution/mousef.html>  
 This genome was contributed by the Brain Molecular Anatomy Project  
 (BMAP)

## FEATURES

Location/Qualifiers  
 1..782  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CS7BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:684199"  
 /tissue\_type="Whole brain"

## source

/dev stage="1, 5, and 15 days newborn"  
 /lab host="DH10B (T1 phage resistant)"  
 /clone lib="NIH\_BMAP\_GH0"  
 /note="Organ: Brain; Vector: pYX-Asc; Site\_1: EcoR I;  
 Site\_2: Not I; the library was constructed according  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction. Ligated  
 with EcoR I adaptor, digested with NotI and then cloned  
 directionally into pYX-Asc vector. The library tsg  
 sequence located between the Not I site and the polyA tail  
 is CGAAGTCAAT. This library was created for the University  
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the  
 Developing Mouse Nervous System', supported by National  
 Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,  
 program coordinator."

## ORIGIN

Alignment Scores:  
 Pred. No.: 5.36e-98 Length: 782  
 Score: 134.00 Matches: 134  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 9.38% Indels: 0  
 DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CB521623 (1-782)

QY 1296 ValAlaAlaValValThrLysIleLysGlnSerCysTrpAlaSerArgArgGluGln 1315  
 DB 3 GTGGCTGCAGTGTCTGGAGATCAACAGAGCTGTGGGCAATCCAGGCGAGAGCAA 62  
 QY 1316 LeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeu 1335  
 DB 63 CTCTCTCGGAGATGCAACAGATGGCGGCGCCCTTTGCTTCTGTAAACGTTGCTTG 122  
 QY 1336 GluThrAspGluProProAspLeuIleGlySerIleLysThrValProLysPro 1355  
 DB 123 GAAACAGATGAAGAACTCTCTGATCTCATTTGGGGGAAGTATAAGACCCCTTCTAAGCCC 182  
 QY 1356 IleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeu 1375  
 DB 183 ATTGCCCTGGAGCCCTGCTTTGGTAACAAAGCCGAGTCTCTCTGTATTCGTGAGGCTC 242  
 QY 1376 ProArgGlyLeuGlyGlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAla 1395  
 DB 243 CCTCGAGGACTGGAGAAATCCCTCTCTGTCAGTCAGGTCCTGCTGGCCAGTGCC 302  
 QY 1396 LeuValAspIleSerGlnGlnMetProIleValTyLysGlyLysSerGlyAlaValArg 1415  
 DB 303 CTGCTGCATTTCTCAGCAGATGCCAATAGTGTACAGGAGAGTCAGGAGCTGTAGA 362  
 QY 1416 AsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429  
 DB 363 AACCGAAGCAGCAGCGCCCTGCACAGCCTGGAACCTGCATT 404

## RESULT 28

LOCUS BF853006 598 bp mRNA linear EST 16-JAN-2001  
 DEFINITION MR2-EN0093-191200-001-b07 EN0093 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF853006  
 VERSION BF853006.1 GI:12240854  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 598)  
 Dias Neto,B., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

COMMENT

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the PAPESP/LICR Human Cancer Genome

project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-

191200-001-b07&t3=2000-12-19&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 598.

Location/Qualifiers

FEATURES

1..598

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="EN0093"

/note="Organ: lung normal; Vector: puc18; Site: 1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application

No. 196,716 - Ludwig Institute for Cancer Research)

profiles into the puc 18 vector. Reverse transcription of

tissue mRNA and cDNA amplification were performed under

low stringency conditions."

ORIGIN

Alignment Scores:

Pred. No.: 2,678-97 Length: 598

Score: 133.00 Matches: 167

Percent Similarity: 98.82% Conservative: 0

Best Local Similarity: 98.82% Mismatches: 1

Query Match: 9.31% Indels: 2

DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BF853006 (1-598)

QY 1095 GlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHis 1114

Db 6 GCGTCTACGGTGTATCCCAATGAGGGAATGTGACCAATGCAAGTCAATGGGCAC 65

QY 1115 AlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGly 1134

Db 66 GCGTCTCTGTGCAACACCAACGCGCAAGTGTCTTGCCACCAACAGGGCGTCAAGGGG 125

QY 1135 AspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThr 1154

Db 126 GACGAGTGCAGCTATGTGAGGTAGTAATTCGATACCAAGGAACCCCTCTCAGAGGAACA 185

QY 1155 CysTyrTyrThrLeuLeuLeuLeuTyrGlnPheThrPheSerLeuSerGlnGluAsp 1174

Db 186 TGTATTATATCTCTCTTATTGACTATCATCTTACCTTTAGTCTATCCAGGAAGATGAT 245

QY 1175 ArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAsp 1194

Db 246 CGGTATTACAGCTATCATTTTGTGGTACTCTCTGACGACAAACAGGATTTGGAC 305

QY 1195 MetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaIleAsnPheSer 1214

Db 306 ATGTTTCATCAATGCTTCCAGAGATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCA 365

QY 1215 AlaGlyThrGlnAlaGlyGluGluMetProValValSerLys-ThrAsnIleLysGlu 1234

Db 366 GTGGACCCAGCGTGGAGAGAGATGCTGTGTTTC-AAGAACCACATTAAAGGAGTA 424

QY 1234 rLysAspSerPheSerAsnGluLysPheAspPheArgHisProAsnIleThrPhePh 1254

Db 425 CAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCAACCAATATCACTTCTT 484

QY 1254 eValTyrValSerAsnPheThrTip 1262

Db 485 TGTTTACGTCAGTAATTTCACCTGG 509

RESULT 29

CA502816/c

LOCUS

DEFINITION

UI-CF-FNO-afg-d-10-0-UI-s1 UI-CF-FNO Homo sapiens cDNA clone

739 bp mRNA linear EST 14-NOV-2002

CA502816

VERSION

CA502816.1 GI:24993770

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 739)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL

MEDLINE

PUBMED

COMMENT

Contact: McCray, PB

McCray Lab

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com) or from Open Biosystems

(www.openbiosystems.com).

The following repetitive elements were found in this cDNA

sequence: 16-237, >L1PA2#LINE/L1 (matched complement)

Seq primer: M13 FORWARD

POLYA=Yes.

FEATURES

Location/Qualifiers

1..739

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="UI-CF-FNO-afg-d-10-0-UI"

/tissue\_type="Human Lung Epithelial cells"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone\_lib="UI-CF-FNO"

/notes="Organ: Lung; Vector: p77T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-CF-FNO is a subtracted cDNA library derived from two

normalized Human lung epithelial cell libraries (EN1 and

DUI) The library was subtracted according to according to

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. For additional information, contact:

bento-soares@uiowa.edu

TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS

chr to LPS 24h

TAG LIB=UI-CF-FNO

TAG\_SEQ=CTGCTCAGGT

ORIGIN

Alignment Scores:

Pred. No.: 5,73e-94 Length: 739

Score: 129.00 Matches: 129  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 9.03% Indels: 0  
 DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CA502816 (1-739)

QY 1139 LeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrThr 1158  
 DB 622 CTATGTGAGGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGACATGTTATTATATCT 563  
 QY 1159 LeuLeuLeuAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrThr 1178  
 DB 562 CTCTCTATTGACATCATCAGTTACCTTTAGTCTATCCAGGAAGATGCGCTATTACACA 503  
 QY 1179 AlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsn 1198  
 DB 502 GCTATCAATTTTGTGGCTACTCTGACGAAACAAACAGGGATTGGACATGTTCAATCAAT 443  
 QY 1199 AlaSerLysAsnPheAsnLeuAsnIleThrTTPAlaAlaSerPheSerAlaGlyThrGln 1218  
 DB 442 GCCTCCAAAGATTTTCAACCTCAACATCACTCGGCTGCCAGTTTCTCAGCTGGACCCAG 383  
 QY 1219 AlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPhe 1238  
 DB 382 GCTGGAGAAGATGCTCTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTC 323  
 QY 1239 SerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSer 1258  
 DB 322 TCTATGAGAGTTTATGATTTTGCACACACCCCAATATCACTTTCTTTGTTATGTCAGT 263  
 QY 1259 AsnPheThrTrpProIleLysIleGln 1267  
 DB 262 AATTTTCACTGGCCCATCAAAATTCAG 236

## RESULT 30

CA895878 503 bp mRNA linear EST 20-DEC-2002  
 B0195B01-5N NIA Mouse Neural Stem Cell (Differentiated) cDNA  
 Library (Long) Mus musculus cDNA clone NIA:B0195B01 IMAGE:30102924  
 5', mRNA sequence.

## ACCESSION

CA895878

## VERSION

CA895878.1 GI:27347427

## KEYWORDS

EST.

## SOURCE

Mus musculus (house mouse)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus; Mus.  
 1 (bases 1 to 503)  
 Piao, Y., Dudekula, D.B., Qian, Y., Martin, P.R., Aiba, K., Vescovi, A.L.  
 and Ko, M.S.H.  
 Systematic Analyses of NIA Mouse Neural Stem Cell (Differentiated)  
 cDNA Library (Long)

## JOURNAL

Unpublished (2002)  
 Contact: Dawood B. Dudekula  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@igsun.grc.nia.nih.gov  
 Plate: B0195 row: B column: 01  
 Seq primer: -21M13 Reverse  
 High quality sequence stop: 503  
 POLYA=No.

## FEATURES

## source

1. 503  
 Location/Qualifiers  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CD1"  
 /db\_xref="niaEST:B0195B01-5N"  
 /db\_xref="taxon:10090"  
 /clone="NIA:B0195B01 IMAGE:30102924"  
 /dev\_stage="Adult"

/lab host="DH10B"  
 /clone\_lib="NIA Mouse Neural Stem Cell (Differentiated)  
 cDNA Library (Long)"  
 /note="vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:  
 NotI; Mouse cDNA project by the Laboratory of Genetics,  
 National Institute on Aging (NIA), Intramural Research  
 Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is  
 a long-transcript enriched cDNA library (Ref. Genome Res.  
 11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were  
 obtained from Dr. Angelo L. Vescovi (Institute for Stem  
 Cell Research, Italy). Double-stranded cDNAs were  
 synthesized with an Oligo(dT) primer [Invitrogen:  
 5'-pGACTAGTTCGTAGATCGGCGCGCCCTTTTCTTTT-3'] from  
 2.0 Microgram of total RNA, treated with T4 DNA  
 polymerase, and purified by ethanol-precipitation. The  
 cDNAs were ligated to lone-linker L1-Sal4, purified by  
 phenol/chloroform, and separated from free linkers by  
 Centricon 100. Then, the cDNAs were amplified by  
 long-range high fidelity PCR using Ex Taq polymerase  
 (Takara) with a primer Sal4-S. The products were purified  
 by phenol/chloroform and Centricon 100. The cDNAs were  
 digested with SalI and NotI enzymes and cloned into  
 SalI/NotI site of pSPORT1 plasmid vector. The DH10B E.  
 coli host was transformed with the ligation mixture by the  
 standard chemical method. The average insert size is about  
 3.2 kb. The library was constructed by Yulan Piao."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.86e-88 Length: 503  
 Score: 122.00 Matches: 122  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 8.54% Indels: 0  
 DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CA895878 (1-503)

QY 1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243  
 DB 138 CCGTGTGTTTCAAAACCAACATCAAGGAATACAAAGATAGCTTCTTAATCAGAAATTT 197  
 QY 1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrPro 1263  
 DB 198 GATTTTCGCAACATCCAAACATCACTTTCTTTGTTTATGTCAGTAATTCATTGGCCC 257  
 QY 1264 IleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePhe 1283  
 DB 258 ATCAAAATTCAGATTGCTTCTCCAGCACAGCAACTTCATGAGCTGGTACAGTTCTTC 317  
 QY 1284 ValThrPhePheSerCysPheLeuSerLeuLeuValAlaValValTrpLysIle 1303  
 DB 318 GTGACTTTCTTCAGTTGTTTCTCTCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 377  
 QY 1304 LysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGlnMet 1323  
 DB 378 AGCAGAGCTGTGGGATCCAGCGGAGAGCAACTTCTTCGGAGATGCAACAGATG 437  
 QY 1324 AlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProAsp 1343  
 DB 438 GCCAGCGCCCTTTGCTTCTGTAACGTTGCTTTGGAACAGATGAAGAACCTCTCAT 497  
 QY 1344 LeuIle 1345  
 DB 498 CTCATT 503

## RESULT 31

CA320165 686 bp mRNA linear EST 09-JUL-2003  
 CA320165  
 UI-M-FW0-cbx-g-11-0-UI\_r1 NIH BMAP\_FW0 Mus musculus cDNA clone  
 DEFINITION  
 IMAGE:6815748 5', mRNA sequence.  
 ACCESSION  
 CA320165  
 VERSION  
 CA320165.1 GI:24538289

```

KEYWORDS      EST.
SOURCE         Mus musculus (house mouse)
ORGANISM       Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE          NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL        National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT        Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-re@mail.nih.gov
               cDNA Library Preparation: Dr. Jim Lin, University of Iowa
               cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
               DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
               Clone Distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/LLNL at:
               http://image.llnl.gov
               This clone was contributed by the Brain Molecular Anatomy Project
               (BMAP)
Seq primer: pYX-5
              Location/Qualifiers
              1..686
              /organism="Mus musculus"
              /mol_type="mRNA"
              /strain="C57BL/6"
              /db_xref="taxon:10090"
              /clone="IMAGE:6815748"
              /tissue_type="whole brain"
              /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
              /lab_host="DH10B (T1 phage resistant)"
              /clone_lib="NIH BMAP FW0"
              /note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I;
              Site 2: Not I; The library was constructed according
              Bonafdo, Lennon and Soares, Genome Research, 6:791-806,
              1996. Denatured RNA was size fractionated on a 1% agarose
              gel. First strand cDNA synthesis was primed with oligo-dT
              primer containing a Not I site. Double strand cDNA was
              size selected according to mRNA size fraction, ligated
              with Ecor I adaptor, digested with NotI and then cloned
              directionally into pYX-Asc vector. The library tag
              sequence located between the Not I site and the polyA tail
              is ACGGACAG. This library was created for the University
              Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
              Developing Mouse Nervous System', supported by National
              Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
              program coordinator."

ORIGIN
Alignment Scores:
Pred. No.:      1.63e-87      Length:      686
Score:          121.00      Matches:     134
Percent Similarity: 99.26%      Conservative: 0
Best Local Similarity: 99.26%      Mismatches: 0
Query Match:      8.47%      Indels:      1
DB:              14      Gaps:        0

US-09-787-097-12 (1-1429) x CA320165 (1-686)
QY 1296 ValAlaAlaValValTrrpLysLleGlnSerCysTrpAlaSerArgArgGluGln 1315
Db 2 GTGGCTGCAGTGTCTGGAGATCAAGCAGAGCTGTGGCATCCAGCGGAGAGCAA 61
QY 1316 LeuLeu-ArgGluMetGlnMetAlaSerArgPropheAlaSerValAsnValAlaLe 1335
Db 62 CTCTCTCCGGGAGATCAACAGATGGCCAGCGCCCTTTGCTCTGTAAACGTGCCTT 121
QY 1335 uGluThrAspGluGluProAspLeuLleGlySerLleLysThrValProLysPr 1355
Db 122 GGAAACAGATGAACACCTCTCTGATCTCATTTGGGGAGAGTATAAGACCGTTCTTAAGCC 181
QY 1355 oileAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLue 1375

```

---

```

Db 182 CATTGCCCTGGAGCCCTCTTTGGTAACAAAGCCAGCTCTCTCTGTATTCTGTAGGCT 241
QY 1375 uProArgGlyLeuGlyGlyLeuProProGlyGlnSerGlyLeuAlaValAlaSerAl 1395
Db 242 CCTCAGGACTGGAGGAATCCCTCTCTCTGTGTAGTCAAGTCTCGCTGTGGCAGTGC 301
QY 1395 aLeuValAspIleSerGlnGlnMetProLleValTrrpLysGluLysSerGlyAlaValAr 1415
Db 302 CCTGTGGACATTCTTCAGCAGATGCCAATAGTGTACAAGGAGAGTCAGGAGCTGAAG 361
QY 1415 gAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db 362 AAACCGAAGCAGCAGCCGCTGCACAGCTGGAACCTGCAIT 404

RESULT: 32
LOCUS      BI089332      706 bp      mRNA      linear      EST 20-JUN-2001
DEFINITION 602832282F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:4994735 5',
            mRNA sequence.
ACCESSION  BI089332
VERSION    BI089332.1 GI:14507662
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 706)
            NIH-MGC http://mgc.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-re@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Incyte Genomics, Inc.
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LNA1017 row: g column: 24
            High quality sequence stop: 676:
FEATURES             Location/Qualifiers
     source            1..706
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:4994735"
                        /cell_line="MGC36"
                        /lab_host="DH10B"
                        /clone_lib="NIH MGC 10"
                        /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
                        Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
                        Average insert size 1.5 kb. Library prepared by Life
                        Technologies."

ORIGIN
Alignment Scores:
Pred. No.:      7.02e-86      Length:      706
Score:          119.00      Matches:     180
Percent Similarity: 98.90%      Conservative: 0
Best Local Similarity: 98.90%      Mismatches: 1
Query Match:      8.33%      Indels:      2
DB:              12      Gaps:        0

US-09-787-097-12 (1-1429) x BI089332 (1-706)
QY 508 AsphHisArgThrArgAlaLeuTrrpValHisGlyClyTrrpLysAlaPheSerAlaAsnLys 527
Db 32 GACCATAGGACGAGCGCCTATACGTTTCATGGTGGCTACAGGCTTTTCAGTGCCTAAG 91
QY 528 TrrArgLeuAlaAspAspLeuTrrArgTrrArgTrrValAspThrGlnMetTrrIleLeu 547
Db 92 TACCGCTTGCAATGATCTCTACCGATATGATGTGGATACCCAGATGTGACCATCTTT 151

```

QY 548 LysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeu 567  
 Db 152 AAGGACAGCCGATTTTCCGTTACTTGGCACACAGCTGTGATAGTGAACCATGCTG 211  
 QY 568 ValPhe-GlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSe 587  
 Db 212 GTGTA-TGGAGGAACACACACATGACACATCTATGAGCCATGGCCCAATGCTTCTC 270  
 QY 587 rSerAspPheMetAlaTyrAspIleAlaCysAspArgTyrSerValLeuProArgProAs 607  
 Db 271 TTCAGATTTTCATGGCCTATGACATTCCTGTGACCGCTGGTCAGTGTCTCCAGACCTGA 330  
 QY 607 PleuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTy 627  
 Db 331 TCTCCACCCATGATGTCAACAGATTTGGCCATTCAGCAGCTTTACACACAGACCATGTA 390  
 QY 627 rValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGluG1 647  
 Db 391 TGTGTTGCGTGGTTTCAATAGTCTCTCTCCTCAGCAGCATCTGGTATTACCTCGGAACA 450  
 QY 647 nCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysVa 667  
 Db 451 GTGTGATGCGCATCGGAGTGAAGCCGCTTGTCTAGCAGCAGGACCTGGTATACGGTGTGT 510  
 QY 667 lTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluG1 687  
 Db 511 GTGGACACAGGGTGTCTCAGTGTATCTCTGTTGGCGCTGGCACTGATGACACAGAGA 570  
 QY 687 uIys 688  
 Db 571 AAAG 574

RESULT 33  
 LOCUS CD644584  
 DEFINITION AGENCOURT\_14539338 NTA Human H1 Embryonic Stem Cell cDNA Library  
 (Long) Homo sapiens cDNA clone IMAGE:30421895 5', mRNA sequence.  
 ACCESSION CD644584  
 VERSION CD644584.1 GI:31816608  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 792)  
 NIH-MGC <http://mgi.nci.nih.gov/>  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: [cgaabs-r@mail.nih.gov](mailto:cgaabs-r@mail.nih.gov)  
 Tissue Procurement: Irene Ginis and Mahendra Rao, NIA  
 cDNA Library Preparation: Yulan Piao and Minoru Ko  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone Distribution: MGC c lone distribution information  
 can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: NDAM500 row: 9 column: 24  
 High quality sequence stop: 681.  
 Location/Qualifiers  
 1..792  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30421895"  
 /tissue\_type="Embryonic Stem cells"  
 /cell\_line="WA01"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_lib="NTA Human H1 Embryonic Stem Cell cDNA Library"

## FEATURES

source

(Long) "  
 /note=vector: pCMV-Sport6; Site 1: NotI; Site 2: SalI;  
 This is a long-transcript enriched cDNA library (Genome  
 Res. 11: 1553-1558 (2001)). [PMID: 11544199] from WA01  
 cell line. Undifferentiated human ES cell line WA01/H1  
 was obtained from WiCell Research Institute, Inc.,  
 Madison, WI, cultured according to their instructions, on  
 MEP feeders. They formed round colonies with defined edges  
 and were positive for alkaline phosphatase, SSEA-4, OCT3,  
 OCT4, REX1, UTF, TERT, SOX2, CX43 and CX45. They are  
 negative for GATA2, GATA4, PDX1, NCAM, MSX1, FLT3, SSEA-1,  
 TUBB3, NES, GFAP, and BOMES. When confluent (18-10 days  
 after plating), the ES cells from 4 X 6cm dishes were  
 treated with 1 mg/ml collagenase, type IV  
 (Invitrogen/GIBCO) for 5-10 min and gently scraped off  
 with 5 ml pipette. RNA was purified with Trizol Reagent  
 from Invitrogen. Protocol ref: Genome Res. 11: 1553-1558  
 (2001). [PMID: 11544199] Double-stranded cDNAs were  
 synthesized with an oligo(dT) primer [Invitrogen:  
 5'-pGAGTAGTCTAGATCGGAGCGCGCCCTTTT-TTTT-3'] from  
 3.4g of total RNA, precipitated with T4 DNA polymerase, and  
 purified by ethanol-precipitation. The cDNAs were ligated  
 to lone-linker LL-Sal4, purified by phenol/chloroform  
 extraction, and separated from free linkers by  
 Centricon-100 column. Then, the cDNAs were amplified by  
 long-range high fidelity PCR using Ex Taq polymerase  
 (Takara) with a primer Sal4-S for 25 cycles. The products  
 were purified by phenol/chloroform extraction and  
 Centricon-100 column. The cDNAs were digested with SalI  
 and NotI enzymes and cloned into SalI/NotI site of  
 pCMV-SPORT6 plasmid vector. The average insert size is  
 about 3.6kb."

## ORIGIN

Alignment Scores:  
 Pred. No.: 7,85e-86 Length: 792  
 Score: 119.00 Matches: 219  
 Percent Similarity: 99.10% Conservative: 0  
 Best Local Similarity: 99.10% Mismatches: 1  
 Query Match: 8.33% Indels: 2  
 DB: 14 Gaps: 0  
 US-09-787-097-12 (1-1429) x CD644584 (1-792)  
 QY 161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180  
 Db 14 CTCATTGAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTCTACAGAG 73  
 QY 181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200  
 Db 74 TGTAGTTGGACCAATTTATATGTTATGATGGGAGCTCAATTTATGACCGGTAGTGT 133  
 QY 201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220  
 Db 134 GCATTTAGTGGCTCATTTGTTCTCTGAGAGAGATGGCAATGAGACTGCTCCCTGAGGTGT 193  
 QY 221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240  
 Db 194 GCCACATCAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 253  
 QY 241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260  
 Db 254 TTTAATATTACTTACAGTTTGTATGTTGTTCCAAATGCTGCTGCTGCTGCTGCTGCTGCTG 313  
 QY 261 LysIleSerAsnSerSerGlu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGly 280  
 Db 314 AAGATCAGTAATAGACGCA-TACTGTTGAATGTGAATGTTCTGAAACTGGAAGGTGA 372  
 QY 280 uAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAs 300  
 Db 373 AGCATGTGACATCTCTCACTGTACAGACAACTGTGTGTTTCTCTCATCAGGAGGATCTCGAA 432  
 QY 300 nSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVa 320

```

Db      433  TTCAAGTGATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAGGTCCTGCGATGTTTCAGT 492
Qy      320  lproValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPr 340
Db      493  TCCGTGTACCAGCTAACCGATCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCC 552
Qy      340  oArgAlaSerHisLysAlaValAlaValAsnGlyAsnIleMetTrpValValGlyGlyTyrMe 360
Db      553  CAGAGCATCTCATAAAGCTGTGGTCAATGGAACATTAATGTGGTGTGGAGGATATAT 612
Qy      360  tPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTTrpLe 380
Db      613  GTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGCT 672
Qy      380  u 380
Db      673  T 673

RESULT 34
BG673763 390 bp mRNA linear EST 04-FEB-2002
LOCUS 9010 ciliated epithelial cDNA cell library Homo sapiens cDNA 5',
DEFINITION mRNA sequence.
ACCESSION BG673763
VERSION BG673763.2 GI:18490081
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 390)
AUTHORS Maiti,A.K., Jorissen,M. and Bouvagnet,P.
TITLE Isolation, in silico characterization and chromosomal localization
of a group of cDNAs from ciliated epithelial cells after in vitro
ciliogenesis
JOURNAL Genome Biol. 2 (7), RESEARCH0026 (2001)
MEDLINE 21407920
PUBMED 11516339
COMMENT On May 1, 2001 this sequence version replaced gi:13919402.
Contact: Maiti AK
Laboratory de Genetic Molculaire Humaine, Faculty de Pharmacy
University Cl. Bernard,
8 Avenue Rockefeller, F69373 Lyon cedex, France
Tel: (33) 478 77 44 25
Fax: (33) 478 77 75 68
Email: amit.maiti@medecine.unige.ch
Seq primer: M13 Forward and reverse.
FEATURES
Location/Qualifiers
source
1..390
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="Nasal biopsy"
/cell_type="cultured ciliated epithelial cells after in
vitro ciliogenesis"
/dev_stage="Adult"
/clove_lib="ciliated epithelial cDNA cell library"
/notes="Vector: PC82.1, cloned RT-PCR fragment from
isolated total RNA"

ORIGIN
Alignment Scores: 1.64e-84 Length: 390
Pred. No.: 117.00 Matches: 117
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 8.19% Indels: 0
Query Match: 12 Gaps: 0
DB:

US-09-787-097-12 (1-1429) x BG673763 (1-390)

Qy      285  ProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVal 304

```

```

Db      2  CTCACTGTACAGACAACCTGGTGTTCCTCATCGAGGCATCTGCAATTCAGTGAATGC 61
Qy      305  ArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAla 324
Db      62  AGAGGATGCTCCTGCTTCTCAGACTGGCAGGTCCTGGATGTTTCAGTTCCTGTACCAGCT 121
Qy      325  AsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHis 344
Db      122  AACCACTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCGAGAGCATCTCAT 181
Qy      345  LysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSer 364
Db      182  AAAGCTGTGGTCAATGGAACATTAATGTGGTGTGGAGGATATATGTTCAACCACTCA 241
Qy      365  AspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArg 384
Db      242  GATTATAACATGCTTCTAGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCGT 301
Qy      385  SerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLys 401
Db      302  TCTGTGAACAAATGCTGTGTAGATATGTCATCTTGTGGCATTATACAG 352

RESULT 35
BE259981 493 bp mRNA linear EST 26-OCT-2000
LOCUS 601148539F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:3163953 5',
DEFINITION mRNA sequence.
ACCESSION BE259981
VERSION BE259981.1 GI:9130910
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 493)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: sgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCMI25 row: m column: 10
High quality sequence stop: 490.
Location/Qualifiers
source
1..493
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3163953"
/tissue_type="neuroblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 19"
/notes="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5',
adaptor: GGCACGAG(G). Library constructed by Ling Hong
in the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores: 1.34e-83 Length: 493
Pred. No.: 116.00 Matches: 149
Score: 98.03% Conservative: 0
Percent Similarity: 98.03% Mismatches: 0
Best Local Similarity: 98.03%

```

```

Query Match:      8.12%      Indels:      3
DB:               10         Gaps:        0
US-09-787-097-12 (1-1429) x BE259981 (1-493)

QY 414 ThrGlyAsnValThrAnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeu 433
DB 41 ACTGGGAATGTGCAATGAGTGTGAGAGTTTTCACATTCAATGAGTCATGGGTGTG 100
QY 434 LeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSer--AlaHisIle-Val 452
DB 101 TTGACCCCTTAAGCAAGAGCAAGTATGAGTGTGGCATCTCCGGCACATGTTGTT 160
QY 453 ThrLeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGly 472
DB 161 ACACCTGAAGAATGGCCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 220
QY 473 TyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThr 492
DB 221 TATATGAAGCAATGTGCAGGAATATGATTTGGATAGAACACATGGAGTATATTACAC 280
QY 493 GlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArg 512
DB 281 CAGGGTGCCTTGTGCAAGGGGTTACGCCATAGCAGTGTTTACGACCATAGGAC 340
QY 513 AlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaSer 532
DB 341 GGCCTATACGTTTCATGGTGGCTACAGGCTTTTCAGTGCCTAATAGTACCGGCT 400
QY 533 AspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhe 552
DB 401 GATCTCTACGATATGATGTGATACCCAGATGTGACCATCTTTTAGGACAGCCG 460
QY 553 PheArgTyrLeuHisThrAlaValLeuVal 562
DB 461 TTCCGTTACTTGCACACAGCTGTGATGTG 490

RESULT 36
BF880955 399 bp mRNA linear EST 17-JAN-2001
LOCUS QV1-ET0181-021200-519-d10 ET0181 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF880955
ACCESSION BF880955.1 GI:12271081
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 399)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,K.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV1&t2=QV1-ET0181-021200-519-d10&t3=2000-12-02&t4=1)
Seq primer: puc 18 forward

```

```

FEATURES
    source          High quality sequence start: 11
                   High quality sequence stop: 399.
    Location/Qualifiers
    1..399
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /dev_stage="Adult"
    /clone_lib="ET0181"
    /note="Organ: lung tumor; Vector: puc18; site 1: SmaI;
    site2: SmaI; A mini-library was made by cloning products
    derived from ORESTES PCR (U.S. Letters Patent application
    No. 196,716 - Ludwig Institute for Cancer Research)
    profiles into the pUC 18 vector. Reverse transcription of
    tissue mRNA and cDNA amplification were performed under
    low stringency conditions."

ORIGIN
Alignment Scores:      7.03e-83      Length:      399
Pred. NO.:            115.00         Matches:     115
Score:                100.00%        Conservative: 0
Percent Similarity:   100.00%        Mismatches:  0
Best Local Similarity: 8.05%         Indels:      0
Query Match:          10             Gaps:        0
DB:

US-09-787-097-12 (1-1429) x BF880955 (1-399)
QY 1054 AsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsn 1073
DB 27  AACTGGTCTTTCATTACCTGTCAGTTCGCCATGCAACGGCCACAGTAATGCAATCAAT 86
QY 1074 GlnSerIleCysGlnLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIle 1093
DB 87  CAGAGCATCTGTGAGAAGTGTGAGAAGCTGCACACAGGCAAGCACTGCGAGACCTGCATA 146
QY 1094 SerGlyPheTyrGlyAspProThrAnGlyGlyLysCysGlnProCysLysCysAsnGly 1113
DB 147 TCTGGCTTCTACCGTATCCCAATGAGGAAATGTCAGCATGCAAGTGCATGG 206
QY 1114 HisAlaSerLeuCysAsnThrAnThrGlyLysCysPheCysThrThrLysGlyValLys 1133
DB 207 CACGGTCTCTGTGCAACCAACACACGGCAAGTCTTCTGCACACCAACCAAGGGCGTCAAG 266
QY 1134 GlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGly 1153
DB 267 GGGGACGAGTGCAGCTATGTGAGGTAGAAATCGATACCAAGGAACCTCTCAGAGGA 326
QY 1154 ThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSer 1168
DB 327 ACATGTTACTATACTCTTCTTATTGACTATCATGTTACCTTTAGT 371

RESULT 37
BF881492/c 681 bp mRNA linear EST 21-FEB-2003
LOCUS UI-CF-EN1-adh-j-10-0-UI-s1 UI-CF-EN1 Homo sapiens cDNA clone
DEFINITION UI-CF-EN1-adh-j-10-0-UI 3', mRNA sequence.
ACCESSION BF881492
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 681)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Ronaldo,M.F., Lennon,G. and Soares,M.B.
normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: McCray, PB
McCray Lab

```





```

QY 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916
Db 122 TCAGAACCCAGTACTCGGGAGTGAAGGTGCAACCTGCATCAACCCATCAATGGTAGT 181
QY 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936
Db 182 GTCTGTGAAGGCTCTCAACACACAGTCTAAGCAGTCCGACACCATGTGCTTGAGG 241
QY 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTyrCysSerAsnMetLys 956
Db 242 ACAGCATGTGGAGATTGCACGCGCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAG 301
QY 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPhe--GlyGlnCysMetGluT 976
Db 302 CAGTGTGTGGATCCCATGCTATGTGGCTCTT-CCCTTTGTGGCCAGTATGGAA 360
QY 976 rPyrThrMetSerThrCysProGluAsnCys-SerGlyTyrCysThrCysSerHis 995
Db 361 GGTATACGATGAGCACCTGCCCTCCCTGAAATGG-TTCAGGCTACTGTACCTGTAGTCAT 419
QY 996 CysLeuGluGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLys 1015
Db 420 TCGTTGGACCAACAGGCTGTGGCTGTGTCTGATCCAGCAATACTGGCAAGGGAAA 479
QY 1016 CysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsn 1035
Db 480 TGCATAGAGGGTTCCTATTAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTACAGGAAT 539
QY 1036 PheTyrProGlnProLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTyr 1055
Db 540 TTCATATCCACGCC-CTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGG 598
QY 1056 SerPheIleHisCysProAlaCysGlnCysAsnGly 1067
Db 599 TCTTTCATTCAGTCCAGCTTGCCATGCAACGGC 634

RESULT 39
AA350293/c
LOCUS EST57512 Infant brain Homo sapiens mRNA linear EST 21-APR-1997
DEFINITION EST57512 Infant brain Homo sapiens cdna 5' end, mRNA sequence.
ACCESSION AA350293
VERSION AA350293.1 GI:2002620
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Adams,M.D., Kerlavage,A.R., Fields,C. and Venter,J.C.
AUTHORS 3,400 expressed sequence tags identify diversity of transcripts
TITLE from human brain
JOURNAL Nat. Genet. 4, 256-267 (1993)
MEDLINE 93364420
PUBMED 8358434
COMMENT Other ESTs: EST57511 THCL168226
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability additional sequence and expression
information related to this EST, please check the TIGR Human Gene
Index (http://www.tigr.org/tdb/hgi.html)
Seq primer: M13 Reverse.
Location/Qualifiers
1. 400
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):151062"
/db_xref="taxon:9606"
/sex="female"

FEATURES
source
1. .671

```

```

/dev_stage="infant"
/clone_lib="Infant brain"
/notes="Organ: brain; Vector: lafmid BA; Site_1: HindIII;
Site_2: NotI"

ORIGIN
Alignment Scores: 1.23e-79 Length: 400
Pred. No.: 111.00 Matches: 111
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 7.77% Gaps: 0
DB: 9

US-09-787-097-12 (1-1429) x AA350293 (1-400)

QY 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940
Db 334 CCTGCAACACACAGTGTAAAGCAGTCCGACACCATGTGCTTGGAGCAGCATGTGA 275
QY 941 AspCysThrSerGlySerSerGluCysMetTyrCysSerAsnMetLysGlnCysValAsp 960
Db 274 GATTGCACCAACGCGNAGCTCTGAGTGCATGTGTGGCAGTATGGAATGGTATACGATGAGC 215
QY 961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrTyrThrMetSer 980
Db 214 TCCAATGCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGC 155
QY 981 ThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnPro 1000
Db 154 ACCTGCCCCCTGAAATTTGTTCCAGCTACTGTACCTGTAGTCATTGCTTGGAGCAACCA 95
QY 1001 GlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
Db 94 GGCTGTGGCTGTGTACTGATCCAGCATATCGCAAGGGAAGATGATAGAGGGTTCC 35

QY 1021 TyrLysGlyProValLysMetProSerGlnAla 1031
Db 34 TATAAAGGACCAAGTGAAGTGCCTTCGCAAGCC 2

RESULT 40
AA1818764/c
LOCUS EST1818764 671 bp mRNA linear EST 07-MAR-2000
DEFINITION will04.x1 NCI CGAP Ut1 Homo sapiens cdna clone IMAGE:2424606 3'
similar to TR:060295 O60295 KIAA0548 PROTEIN ;contains L1.t3 L1
repetitive element ;, mRNA sequence.
ACCESSION AA1818764
VERSION AA1818764.1 GI:5437843
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 671)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL Tumor Gene Index
COMMENT Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
CDNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 2162 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 411.
Location/Qualifiers
1. .671

```

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:244606"  
/tissue\_type="well-differentiated endometrial  
adenocarcinoma, 7 pooled tumors"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Uti"  
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;  
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
Average insert size 1.75 Kb. Life Technologies catalog #: 11538-014"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,32e-78 Length: 671  
Score: 110.00 Matches: 110  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 7.70% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AI818764 (1-671)

QY 1158 ThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyr 1177  
DB 547 ACTCTTCTTATTGACTATCAGTTCACCTTTCAGTCTATCCCGAAGATGATCGCTATTAC 488  
QY 1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197  
DB 487 ACAGTATCAATTTTGTGGTACTCTCGACGACAAACACAGGATTTGGACATGTTTCATC 428  
QY 1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThr 1217  
DB 427 AATGCTCCACGAATTTCAACTCAACATCACCTGGCTGCGAGTTTCTCAGCTCGAAC 368  
QY 1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237  
DB 367 CAGGCTGGAGAGAGATGCTGTTGTTTCAAAAACCAACATTAAAGGATCAAAAGATAGT 308  
QY 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257  
DB 307 TTCTCTAATGAGAAGTTTGATTTTCGCAACCAACCAATATACATTTCTTTTATGTC 248  
QY 1258 SerAsnPheThrTrpProIleLysIleGln 1267  
DB 247 AGTAATTTCACTGGCCCATCAAAATTCAG 218

## RESULT 41

AL048842  
LOCUS DKF2p434G2118\_r1 434 (synonym: htes3) Homo sapiens cDNA clone  
DEFINITION DKF2p434G2118, mRNA sequence.

## ACCESSION

AL048842

## VERSION

AL048842.1 GI:4728151

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

## REFERENCE

1 (bases 1 to 502)

## AUTHORS

Wiemann, S.

## TITLE

EST (Ottewaelder, et al.)

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: MIPS

## MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

## This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

## Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;

sequenced by MedGenomix (Martinsried/Germany) within the cDNA

available.

This clone (DKF2p434G2118) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

## FEATURES

## source

1..502  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKF2p434G2118"  
/tissue\_type="testis"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="434 (synonym: htes3)"  
/note="Vector: pSPORT1; Site 1: NotI; Site 2: SalI"

## ORIGIN

Alignment Scores:  
Pred. No.: 4.16e-77 Length: 502  
Score: 108.00 Matches: 108  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 7.56% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AL048842 (1-502)

QY 159 ThrTrpLeuIleGluGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAla 178  
DB 2 ACGTGGCTCATTCAGACGACGACCAATAGAAATAGAGACTTCGTTTCAATCATTTTGGCT 61  
QY 179 ThrGluCysSerTrpAspHisIleTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198  
DB 62 ACAGAGTGTAGTTGGGACCAATTATATGTTATGATGGGACTCAATTTATGACCGCTA 121  
QY 199 ValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGlu 218  
DB 122 GTTGTGCTCATTTAGTGGCTCATTTGTTCTCGAGAGATGGCAATGAGACTGTCCCTGAG 181  
QY 219 ValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaIaTyrAsnLeu 238  
DB 182 GTTGTGGCCACATCAGCTTATGCTTGTGTCATTTTTTTAGTGTATGCTGCTTATATTTG 241  
QY 239 ThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGly 258  
DB 242 ACTGGATTTAATATTACTTACAGTTTGTATGATGTGTCCTCAATATCTGCTCAGGCCGAGGA 301  
QY 259 GluCysLysIleSerAsnSerSer 266  
DB 302 GAGTGTAAAGATCAGTAATAGCAGC 325

## RESULT 42

BF963698/c

## LOCUS

PM4-NN1204-221200-001-g03 NN1204 Homo sapiens cDNA, mRNA sequence.

## ACCESSION

BF963698

## VERSION

BF963698.1 GI:12380973

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Homo sapiens

## REFERENCE

1 (bases 1 to 484)

## AUTHORS

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,

O'Hare, M.J.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

## TITLE

Shotgun sequencing of the human transcriptome with ORF expressed

## JOURNAL

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

## MEDLINE

20202663

**PUBMED** 10737800  
**COMMENT** Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM4tt2-PM4-NN1204-221200-001-903&ts=2000-12-22&f=1>)  
 Seq primer: puc 18 forward  
 High quality sequence start: 8  
 High quality sequence stop: 350.  
**FEATURES** Location/Qualifiers  
 1..484  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="NN1204"  
 /note="Organ: nervous normal; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."  
**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 7,038-74 Length: 484  
 Score: 104.00 Matches: 135  
 Percent Similarity: 98.54% Conservative: 0  
 Best Local Similarity: 98.54% Mismatches: 1  
 Query Match: 7.28% Indels: 2  
 DB: 10 Gaps: 0  
 US-09-787-097-12 (1-1429) x BF963698 (1-484)  
 QY 1025 VallysMetProSer:GlnAlaProThrGlyAsnPhetYrProGlnProLeuLeuAsnSer 1044  
 DB 482 GTGAAGATGCTTCGACAGCCCTACAGAAATTTCTATCACAGCCCTGCTCAATTCC 423  
 QY 1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGln 1064  
 DB 422 AGCATGTGCTAGAGGACAGACAGATACAACTGGTCTTTTCATCTACTGTCCAGCTTGCCAA 363  
 QY 1065 CysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThr 1084  
 DB 362 TCGAAGCGCCACAGTAATGATCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGACC 303  
 QY 1085 ThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGly 1104  
 DB 302 ACAGCAAGACACTGGAGACCTGCATATCTGGCTTCTACGGTGATCCCAATGGAGGG 243  
 QY 1105 LysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLys 1124  
 DB 242 AAATCTCAGCCATGCAAGTGCATGAGGACCGGCTCTGTGCACACCAACACGGGCAAG 183  
 QY 1125 CysPheCysThrThr-LysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAs 1144  
 DB 182 TGCCTTCGACCNC-CAAGGGCGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAA 124  
 QY 1144 nArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160  
 DB 123 TCGATACCAAGGAACCTCTTCAGAGGAACATGTTATTACTCTTCT 75  
**RESULT 43**  
 AI372809/c  
 LOCUS 383 bp mRNA linear EST 12-JAN-1999

**DEFINITION** EST175353 Infant brain, Bento Soares Homo sapiens cDNA clone  
**FEATURES** FUHIBU1, mRNA sequence.  
**ACCESSION** AI372809  
**VERSION** AI372809.1 GI:4152675  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
**REFERENCE** 1 (bases 1 to 383)  
**AUTHORS** Adams, M.D., Kexlavage, A.R., Fleischmann, R.D., Fuldner, R.A.,  
 Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstein, K.G., Gocayne, J.D.,  
 White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C.,  
 Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D.,  
 Fitzgerald, D.M., Fitzhugh, W.M., Fritchman, J.B., Geoghagen, N.S.,  
 Glodok, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr.,  
 Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M.,  
 Moreno-Palauques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M.,  
 Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R.,  
 Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y.,  
 Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J.,  
 Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A.,  
 He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K.,  
 Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H.,  
 Raymond, L., Wei, Y.P., Wing, J., Xu, C., Yu, G.L., Ruben, S.M.,  
 Dillion, P.J., Pannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C.,  
 Fraser, C.M. and Venter, J.C.  
**TITLE** Initial assessment of human gene diversity and expression patterns  
 based upon 83 million nucleotides of cDNA sequence  
**JOURNAL** Nature 377 (6547 Suppl), 3-174 (1995)  
**MEDLINE** 96026280  
**PUBMED** 7566098  
**COMMENT** Other ESTs: EST175328 EST175329 EST175330 EST175331 EST175332  
 EST175333 EST175334 EST175335 EST175336 EST175337 EST175338  
 EST175339 EST175340  
 Contact: HGI (Human Gene Index)  
 The Institute for Genomic Research  
 9712, Medical Center Drive, Rockville, MD 20850, USA  
 Tel: (301)-838-0200  
 Fax: (301)-838-0208  
 Email: hgi@igicr.org.  
**FEATURES** Location/Qualifiers  
 1..383  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="FUHIBU1"  
 /lab\_host="E. coli DH5-alpha"  
 /clone\_lib="Infant brain, Bento Soares"  
 /note="Vector: BA, M13-derived; Site 1: HindIII; Site 2:  
 NotI; The infant brain library, constructed by Bento  
 Soares, Columbia University, was oligo-(dT) primed and  
 directionally cloned into an M13-derived plasmid using  
 total brain mRNA from a 72-day old human female afflicted  
 with spinal muscular atrophy."  
**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 6,338-70 Length: 383  
 Score: 99.00 Matches: 99  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.93% Indels: 0  
 DB: 9 Gaps: 0  
 US-09-787-097-12 (1-1429) x AI372809 (1-383)  
 QY 822 LeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIle 841  
 DB 297 CTTTGGCTCTCTTACAAACCCAGAGAGGAGGATTTGCTCTTACGACGTCGCAATA 238  
 QY 842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTyrValGlyLeuArgLys 861

Db 237 ATGAGTCATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGTGGCTTCGGAAG 178  
Qy 862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881  
Db 177 ATCAATGTGTCTACTTGGTGTGGAGATATGCCCCATTACAAATAGTTTACTACAG 118  
Qy 882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901  
Db 117 TGGATGCCCTTGAGCCAGATGCTGGATCTTGTGGAATTTTATCAGAACCCAGTACT 58  
Qy 902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920  
Db 57 CGGGAGCTGAAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGTGAAGG 1

## RESULT 44

LOCUS CA874419 567 bp mRNA linear EST 20-DEC-2002  
DEFINITION K0932C03-SN NIA Mouse Neural Stem Cell (Undifferentiated) cDNA  
Library (long) Mus musculus cDNA clone NIA:K0932C03 IMAGE:30087290  
5', mRNA sequence.

## ACCESSION

VERSION CA874419.1 GI:27325968

## KEYWORDS

SOURCE Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 (bases 1 to 567)  
Piao, Y., Dudekula, D.B., Qian, Y., Martin, P.R., Aiba, K., Vescovi, A.L.  
and Ko, M.S.H.  
Systematic Analyses of NIA Mouse Neural Stem Cell  
(Undifferentiated) cDNA Library (long)  
Unpublished (2002)

## JOURNAL

COMMENT Contact: Dawood B. Dudekula

National Institute on Aging/National Institutes of Health

333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA

Email: cdna@ngs.grc.nia.nih.gov

Plate: K0932 row: C column: 03

Seq primer: -21M3 Reverse

High quality sequence stop: 567

POLYA=No.

## FEATURES

source

1..567

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="CD1"

/db\_xref="taxon:10090"

/db\_xref="taxon:10090"

/clone="NIA:K0932C03 IMAGE:30087290"

/dev\_stage="Adult"

/lab\_host="DH10B"

/clone\_lib="NIA Mouse Neural Stem Cell (Undifferentiated)

cDNA Library (long)"

/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:

NotI; Mouse cDNA project by the Laboratory of Genetics,

National Institute on Aging (NIA), Intramural Research

Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is

a long-transcript enriched cDNA library (Ref. Genome Res.

11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were

obtained from Dr. Angelo L. Vescovi (Institute for Stem

Cell Research, Italy). Double-stranded cDNAs were

synthesized with an Oligo(dT) primer [Invitrogen:

5'-pGACTAGTCTAGATCGAGGCGGCCCTTTTCTTTT-3'] from

2.0 Microgram of total RNA, treated with T4 DNA

polymerase, and purified by ethanol-precipitation. The

cDNAs were ligated to Loxe-linker IL-Sal4, purified by

phenol/chloroform, and separated from free linkers by

Centricon 100. Then, the cDNAs were amplified by

long-range high fidelity PCR using Ex Taq polymerase

(Takara) with a primer Sal4-S. The products were purified

by phenol/chloroform and Centricon 100. The cDNAs were

digested with SalI and NotI enzymes and cloned into

## ORIGIN

Alignment Scores:  
Pred. No.: 9,28e-70 Length: 567  
Score: 99.00 Matches: 99  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.93% Indels: 0  
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CA874419 (1-567)

Qy 842 MetClnSerSerClnSerMetSerLysLeuThrProTrpValGlyLeuArgLys 861  
Db 240 ATGCAATCATCTCAAGTATGTCAGCTCCTCTGCTCCATTCAGTGGTTCGGAAG 299  
Qy 862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881  
Db 300 ATCAATGTGTCTTACTTGGTGTGGAGGATATGCTCCATTCAAAATAGTTTGTCTG 359  
Qy 882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901  
Db 360 TGGATGCCATCTCAGCCAGTGAATGCTTCTGTGGATCTTGTACAGAGCTTAGTACT 419  
Qy 902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
Db 420 CGGGATTAAGGCTGCACCTGCATCAACCTCTCAATGGCAGCGTCTGTGAAGGCT 479  
Qy 922 AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940  
Db 480 GCAAAACACAGTCCCAAGCAGTCCCGACACCATGTGCCCTGCGGACAGCGTGTGGC 536

## RESULT 45

LOCUS BF313142

DEFINITION BF313142.1 NIH\_MGC\_19 Homo sapiens cDNA clone IMAGE:4126959 5',

mRNA sequence.

ACCESSION BF313142

VERSION BF313142.1 GI:11261090

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 953)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

CONTACT: Robert Strausberg, Ph.D.

Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov

Plate: L10M1017 row: j column: 16

High quality sequence stop: 612.

Location/Qualifiers

1..953

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4126959"

/tissue types="neuroblastoma"

/lab host="DH10B (phage-resistant)"

/clone lib="NIH MGC 19"

/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:

ECOR1; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies).  
Note: this is a NIH\_MGC Library."

## ORIGIN

Alignment Scores:  
Pred. No.: 1,54e-69 Length: 953  
Score: 99.00 Matches: 154  
Percent Similarity: 99.35% Conservative: 0  
Best Local Similarity: 99.35% Mismatches: 1  
Query Match: 6.93% Indels: 1  
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BF313142 (1-953)

QY 414 ThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeu 433  
Db 61 ACTGGAAATGTGACCAATGATGAGATTTTTCATTCATATGATGATGATGATG 120  
QY 434 LeuThrProLysAlaLysGluGlnTrpAlaValValGlyHisSerAlaHisIleValThr 453  
Db 121 TTGACCCCTAAGGCAAGGAGCAGTATGAGTGGTGGCACTCTGCACACATTTTACA 180  
QY 454 LeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTrpGlyTyr 473  
Db 181 CTGAAGAATGGCGGAGTGTGATGCTGTCATCTTTGGTCACTGCTCTATGGATAT 240  
QY 474 IleSerAsnValGlnGluTrpAspLeuAspLysAsnThrTrpSerIleLeuHisThrGln 493  
Db 241 ATAAGCAATGTGACGAATATGATTTGGATAGACACATGAGATATATACACCCAG 300  
QY 494 GlyAlaLeuValGlnGlyGlyTyrGlyHisSerValTyrAspHisArgThrArgAla 513  
Db 301 GGTGCCCTTGTGCAAGGGGGTTACGCCATAGCAGTGTGTACGACCATAGGACCCAG-GCC 359  
QY 514 LeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAsp 533  
Db 360 CTATACGTTTCATGTTGGTGTACAGGCTTTCAGTGCATATAGTACCGCTTGCAGATGAT 419  
QY 534 LeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePhe 553  
Db 420 CTCCTACCGATATGATGATGATACCCAGATGTGGACCAATCTTAAAGACACGCCGATTTTC 479  
QY 554 ArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuVal 568  
Db 480 CGTTACTTGCACACAGCTGTGATGATGAGTGAACCATGCTGGTG 524

RESULT 46  
LOCUS BF853681  
DEFINITION MR2-EN0093-211200-003-a05 EN0093 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF853681  
VERSION BF853681.1 GI:12241425  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 595)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongsneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
PUBMED 10737800

## COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=MR2&t2=MR2-EN0093-211200-003-a05&t3=2000-12-21&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 595.

FEATURES  
source

1..595  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="EN0093"  
/notes="Organ: lung normal; Vector: puc18; Site: 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Alignment Scores:  
Pred. No.: 6,29e-69 Length: 595  
Score: 98.00 Matches: 157  
Percent Similarity: 98.74% Conservative: 0  
Best Local Similarity: 98.74% Mismatches: 1  
Query Match: 6.86% Indels: 2  
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BF853681 (1-595)

QY 1095 GlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHis 1114  
Db 7 GGCTTCTACGGTGATCCCAATGAGGGAATGTCACCATGCAAGTGCATGGGCAC 66  
QY 1115 AlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGly 1134  
Db 67 GCCTCTCTGTGCACACACACACGCGGCAAGTCTCTGCACCAAGGCGTCAGGGG 126  
QY 1135 AspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThr 1154  
Db 127 GACGAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAACCCCTCTCAGAGAAC 186  
QY 1155 CysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAsp 1174  
Db 187 TGTATTATATCTCTCTTATTGACTATCAGTTACCTTTAGTCTATCCACGAGATGAT 246  
QY 1175 ArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeu-As 1194  
Db 247 CGCTATTACAGCTATCAATTTTGTGGTACTCTCTGCAGCAACAAACAGGGA-TTGGGA 305  
QY 1194 pMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSe 1214  
Db 306 CATGTTCAATCATGCTCCCAAGATTTCACTCAACATCAGCTGGGTGCCAGTTTCTC 365  
QY 1214 rAlaGlyThrGlnAlaGlyGluMetProValValSerLysThrAsnIleLysGluTyr 1234  
Db 366 AGCTGAACCCAGGCTGAGAGACAGATGCTGTGTTTCAAAAACCAACATTAGAGTA 425  
QY 1234 rLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThr 1252  
Db 426 CAAAGATAGTTTCTCTAATGAGAGTTTGAATTTTCGAACCAACCAACCAATATCACT 480

RESULT 47

```

CB525640
LOCUS      CB525640          565 bp      mRNA      linear      EST 09-JUL-2003
DEFINITION UI-M-FY0-cfd-j-16-0-UI_r1 NIH_BMAP_FY0 Mus musculus cDNA clone
IMAGE:6848081 5', mRNA sequence.
ACCESSION  CB525640
VERSION     CB525640.1  GI:29359097
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
REFERENCE   1  (bases 1 to 565)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-remail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found at
            http://genome.uioawa.edu/distribution/mousefl.html
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
            Seq primer: pYX-5.
FEATURES             Location/Qualifiers
     source           1..565
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="CS7BL/6"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6848081"
                     /tissue_type="whole brain"
                     /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
                     /lab_host="DH10B (TI phage resistant)"
                     /clone_lib="NIH_BMAP_FY0"
                     /note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
                     Site 2: Not I; The library was constructed according
                     Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
                     1996. Denatured RNA was size fractionated on a 1% agarose
                     gel. First strand cDNA synthesis was primed with oligo-dT
                     primer containing a Not I site. Double strand cDNA was
                     size selected according to mRNA size fraction, ligated
                     with EcoR I adaptor, digested with NotI and then cloned
                     directionally into pYX-Asc vector. The library tag
                     sequence located between the Not I site and the polyA tail
                     is ACGGACAG. This library was created for the University
                     Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
                     Developing Mouse Nervous System', supported by National
                     Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
                     program coordinator."
ORIGIN
Alignment Scores:
Pred. No.:      1..628-66      Length:      565
Score:          95.00      Matches:      95
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     6.65%      Indels:    0
DB:              14      Gaps:        0

US-09-787-097-12 (1-1429) x CB525640 (1-565)

QY      1335  LeuGluThrAspGluGluProProAspLeuIleGlySerIleValProIlys 1354
Db      2  TTGGAACAGATGAAGACCTCTGATCTCATTTGGGGAAGTATAAGACCGTCTCTTAAG 61
QY      1355  ProfileAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArg 1374
Db      62  CCCATGCCCTGGAGCCCTGCTTTGGTACAAAGCGCGAGTCTCTGTTATCTGAGG 121
QY      1375  LeuProArgGlyLeuGlyGlyIleProProGlyGlnSerGlyLeuAlaValAlaSer 1394

```

```

Db      122  CTCCTCGAGGACTGGAGGAATCCCTCTCTCTGCTGCTCAGGTCCTGCTGTGGCCAGT 181
QY      1395  AlaLeuValAspIleSerGlnGlnMetProIleValTyIysGluLysSerGlyAlaVal 1414
Db      182  GCCTCTGTGGACATTTCTCAGCAGATGCCAATAGTGTACAAAGAGAGTCAAGAGCTGTA 241
QY      1415  AtgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db      242  AGAAACCGAAGCAGAGCGCGCTGCAGAGCTGGAACTGCATT 286

RESULT 48
LOCUS    AI372810/c
DEFINITION EST1753354 Infant brain, Bento Soares Homo sapiens cDNA clone
FUTIBUL, mRNA sequence.
ACCESSION AI372810  GI:4152676
VERSION    AI372810
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1  (bases 1 to 320)
AUTHORS    Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
            Bult,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
            White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-Wai,C.,
            Clayton,R.A., Cline,T.R., Cotton,M.D., Earle-Hughes,J., Fine,L.D.,
            Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghagen,N.S.,
            Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S.Jr.,
            Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
            Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
            Phillips,C.A., Ryder,S.E., Scott,J.L., Saudek,D.M., Shirley,R.,
            Small,K.V., Spriggs,T.A., Utterback,T.R., Weidman,J.F., Li,Y.,
            Bednarek,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,E.J.,
            Dimke,D., Feng,D.-F., Ferrie,A., Fischer,C., Hastings,G.A.,
            He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
            Kozak,D.D., Kunsch,C., HungJun,J., Li,H., Weissner,P.S., Olsen,H.,
            Raymond,L., Wei,Y.F., Wing,J., Xu,C., Yu,G.L., Ruben,S.M.,
            Dillion,P.J., Fannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
            Fraser,C.M. and Venter,J.C.
            Initial assessment of human gene diversity and expression patterns
            based upon 83 million nucleotides of cDNA sequence
            Nature 377 (6547 Suppl), 3-174 (1995)
            96026280
            7566098
COMMENT    Other ESTs: EST175328 EST175329 EST175330 EST175331 EST175332
            EST175333 EST175334 EST175335 EST175336 EST175337 EST175338
            EST175339 EST17534
Contact: HGI (Human Gene Index)
The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-0200
Fax: (301)-838-0208
Email: hgi@tigr.org.
FEATURES             Location/Qualifiers
     source           1..320
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="FUTIBUL"
                     /lab_host="E. coli DH5-alpha"
                     /clone_lib="Infant brain, Bento Soares"
                     /note="Vector: BA, M13-derived; Site 1: HindIII; Site 2:
                     NotI; The infant brain library, constructed by Bento
                     Soares, Columbia University, was oligo-(dT) primed and
                     directionally cloned into an M13-derived plasmid using
                     total brain mRNA from a 72-day old human female afflicted
                     with spinal muscular atrophy."
ORIGIN
Alignment Scores:
Pred. No.:      3..886-65      Length:      320

```





Pred. No.: 5.98e-62 Length: 281  
 Score: 89.00 Matches: 89  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.23% Indels: 0  
 DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BG015062 (1-281)

QY 588 SerAspPheMetAlaTyAspIleAlaCysAspArgTrpSerValLeuProArgProAsp 607  
 Db 279 TCAGATTTCATGGCTATGACATTGCTGTCAGCGCTGGTCAGTCTCCAGACCTGAT 220  
 QY 608 LeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyr 627  
 Db 219 CTCACCATGATGCAACAGATTGGCCATTACAGCTTTACACAGCACCATTGAT 160  
 QY 628 ValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPheThrSerGluGln 647  
 Db 159 GTGTTGGTGTTCATATGATCTCTCTCAGCGACATCTCGTATTACCTCGGAACAG 100  
 QY 648 CysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysVal 667  
 Db 99 TGTGATGGCATCGAGTGAAGCCGCTGTGTAGCAGCAGGACCTGGTATTCGGTGTG 40  
 QY 668 TrpAsnThrGlySerSerGlnCysIle 676  
 Db 39 TCGAACACAGGGTGTCTCAGTGTATC 13

RESULT 51  
 BF853021/c 483 bp mRNA linear EST 16-JAN-2001  
 LOCUS MR2-EN0093-191200-001-905 EN0093 Homo sapiens CDNA, mRNA sequence.

DEFINITION BF853021

ACCESSION BF853021.1 GI:12240869

VERSION EST.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 483)

Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.P.,

Goldman,G.H., Carvalho,A.F., Matekuma,A., Baia,G.S., Simpson,D.H.,

Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,

O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

CONTACT: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-EN0093-

191200-001-905&t3=2000-12-19&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 483.

Location/Qualifiers

1. 483

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_lib="EN0093"

FEATURES

source

# ORIGIN

Alignment Scores:  
 Pred. No.: 1.02e-61 Length: 483  
 Score: 89.00 Matches: 89  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.23% Indels: 0  
 DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BF853021 (1-483)

QY 1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCystTyrThrLeuLeu 1160  
 Db 446 GAGGTAGAAAATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATTATCTCTT 387  
 QY 1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrThrAlaIle 1180  
 Db 386 ATTGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGCGCTATTACACAGTATC 327  
 QY 1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200  
 Db 326 AATTTTGTGGCTACTCTCTGACGACAAACACAGGATTTGGACATGTTCAATGCTCC 267  
 QY 1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220  
 Db 266 AAGATTTCACCTCAACATCACCTCGGCTGCGCTTCTCAGCTGGAACCCAGGCTGGA 207  
 QY 1221 GluGluMetProValValSerLysThr 1229  
 Db 206 GAAGAGATGCTGCTGTTTTCAAAAC 180

## RESULT 52

AA350292

LOCUS

DEFINITION

AA350292.1

ACCESSION

AA350292.1

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Other\_ESTs: EST57512.JHC168885

Contact: Kerlavage, AR

Bioinformatics

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850 USA

Tel: 3018699056

Fax: 3018699423

Email: arkerlav@tigr.org

For clone availability, additional sequence and expression

information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)

Seq primer: M13-21.

Location/Qualifiers

1. 396

/organism="Homo sapiens"

/mol\_type="mRNA"

FEATURES

source

/note="Organ: lung normal; Vector: puc18; Site 1: SmaI;  
 Site 2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

```

/db_xref="ATCC (inhost):151062"
/db_xref="taxon:9606"
/sex="female"
/dev_stage="infant"
/clone_lib="infant brain"
/note="Organ: brain; Vector: lafmid BA; Site_1: HindIII; Site_2: NotI"

ORIGIN
Alignment Scores:
Pred. No.: 2,266-59 Length: 396
Score: 86.00 Matches: 123
Percent Similarity: 98.40% Conservative: 0
Best Local Similarity: 98.40% Mismatches: 1
Query Match: 6.02% Indels: 2
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AA350292 (1-396)
QY 480 TTAATTCGATAGACACATGGAGTATATACACCCAGGCGCCCTTGCGACGG 60
DB 1 TATGATTCGATAGACACATGGAGTATATACACCCAGGCGCCCTTGCGACGG 60
QY 500 GYTYRGLYHISERVALTYRASPISARGHTRARGALALEUTYRVALHIS-GLYGL 519
DB 61 GGTACCGCCATAGCAGTGTTCAGCACCATAGCAGCCAGGCCCTATACGTTNA-TGGTGG 119
QY 519 YTYRISALAPHESERALASLYSTRARGLEULALASPHLEUTYRARGTYRASPVA 539
DB 120 CTACAAAGGCTTTTCAGTGCCAAATAGTACCGGCTTCAGATGATCTCTACCGATATGATGT 179
QY 539 LASPThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAl 559
DB 180 GGATACCCAGATGTGGACCATCTTAAAGACACGCCGATTTTCCTTACTTGACACAGC 239
QY 559 avallleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMe 579
DB 240 TGTGATAGTACGTGAACCATCTCGTGTGTGGAGGAACACACACATGACATCTTAT 299
QY 579 tSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspAr 599
DB 300 GAGCCATGGCGCAATAGCTTCTTCAGATTTCATGCCCTATGACATGTGCTGTGACCG 359
QY 599 gTrpSerValLeu 603
DB 360 CTGGTCAGTGCCT 372

RESULT 53
CB522182
LOCUS
DEFINITION
UI-M-GH0-cep-c-09-0-UI-r1 NIH_BMAP_GH0 Mus musculus cDNA clone
IMAGE:6842530 5', mRNA sequence.
CB522182
VERSION
CB522182.1 GI:29355537
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefi.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

```

```

Seq primer: pyX-5.
Location/Qualifiers
1. 758
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6842530"
/tissue_type="Whole brain"
/dev_stage="1, 5, and 15 days newborn"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GH0"
/note="Organ: Brain; Vector: pyX-Asc; Site_1: EcoR I; Site_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAATCGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
Alignment Scores:
Pred. No.: 1,786-57 Length: 758
Score: 84.00 Matches: 84
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.88% Indels: 0
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CB522182 (1-758)
QY 1346 GlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLys 1365
DB 15 GGGGGAGTATTAAGACCCCTTCTTAGCCCATTTGCCCTGGAGCCCTGCTTTGGTAACAA 74
QY 1366 AlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyLysLeuProPro 1385
DB 75 GCGCAGTCTCTCTCTATTCTGAGGCTCCCTCGAGGACTGGAGGAATCCCTCTCTCT 134
QY 1386 GlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIle 1405
DB 135 GGTCACTCAGGCTCTCTGTGCCAGTGCCTGGTGACATTTCTCAGCAGATGCCAATA 194
QY 1406 ValTyrLysGlyLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnPro 1425
DB 195 GTGTACAGGAGAGTCAAGAGCTGTATAGAAACCGAAGCAGCAGCGGCTGCACAGCT 254
QY 1426 GlyThrCysIle 1429
DB 255 GGAACCTGCATT 266

RESULT 54
CB531905
LOCUS
DEFINITION
AGENCOURT 10205588 NIH_MGC_126 Homo sapiens cDNA clone
IMAGE:6556819 5', mRNA sequence.
CB531905
VERSION
CB531905.1 GI:22842346
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/

```

**TITLE** National Institutes of Health, Mammalian Gene Collection (MGC)  
**JOURNAL** Unpublished (1999)  
**COMMENT** Contact: Robert Straubeberg, Ph.D.  
 Email: c9pbbs-remail.nih.gov  
 Tissue Procurement: NCI  
 cDNA Library Preparation: Michael Brownstein Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Plate: LICM2723 row: n column: 19  
 High quality sequence stop: 480.  
**FEATURES** Location/Qualifiers  
 1..778  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6556819"  
 /tissue\_type="mixed (pool of 40 RNAs)"  
 /lab\_host="DH10B (TI-phage-resistant)"  
 /clone\_lib="NIH MGC 126"  
 /note="Vector: pDNR-LIB; Site 1: sfii (ggcattatggcc); Site 2: SfiI (ggcgctcgcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:  
 5'-AGCAGTGGTATACGACAGTGGCATTACGCGCGG-3' and 5'-ATTGAGCGCGGAGCGCGCATG-dt(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-1 kb size fraction (other fractions present in NIH MGC 127 and NIH MGC 128). Library created in the laboratory of T. Usdin, M.D., Ph.D. (NIH, NIH). Note: this is a NIH\_MGC Library."  
**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 1,838-57 Length: 778  
 Score: 84.00 Matches: 84  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.88% Indels: 0  
 DB: 13 Gaps: 0  
 US-09-787-097-12 (1-1429) x BUS31905 (1-778)  
 QY 1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPhe 1203  
 DB 4 GCTACTCTCTGACGACAAACAGGAGTTTGACATGTTTCATCAATGCTCCCAAGATTTC 63  
 QY 1204 AsnLeuAsnIleThrTrrAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluMet 1223  
 DB 64 AACCTCAACATCACTGGGCTGCCAGTTCTCAGCTGGAACCCAGGCTGGAGAGAGATG 123  
 QY 1224 ProValValSerLysThrAsnIleLysGluTyrlYrLysAspSerPheSerAsnGluLysPhe 1243  
 DB 124 CCTGTTGTTTCAAAACCAACATTAAAGGAGTACAAAGATAGTTTCTCTTAATGAGAAGTTT 183  
 QY 1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrlValSerAsnPheThrTrrPro 1263  
 DB 184 GATTTCGCAACACCCCAATATCACTTCTTTGTTTATGTCAGTAATTTTCACTGGCCCC 243  
 QY 1264 IleLysIleGln 1267  
 DB 244 ATCAAAATTCAG 255  
**RESULT 55**  
**BI004303**  
**LOCUS** BI004303 484 bp mRNA linear EST 13-JUN-2001  
**DEFINITION** PMO-HN0078-280201-010-h11 HN0078 Homo sapiens cDNA, mRNA sequence.  
**ACCESSION** BI004303  
**VERSION** BI004303.1 GI:14408377  
**KEYWORDS** EST.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
**REFERENCE** 1 (bases 1 to 484)  
**AUTHORS** Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.P., Matsukuma,A., Bais,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Congenel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
**TITLE** Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
**MEDLINE** 20202663  
**PUBMED** 10737800  
**COMMENT** Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PMO&t2=PMO-HN0078-280201-010-h11&t3=2001-02-28&t4=1>)  
 Seq primer: puc 18 forward  
 High quality sequence start: 18  
 High quality sequence stop: 483.  
**FEATURES** Location/Qualifiers  
 1..484  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HN0078"  
 /note="Organ: head normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from CRESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
**ORIGIN**  
 Alignment Scores:  
 Pred. No.: 4,81e-56 Length: 484  
 Score: 82.00 Matches: 124  
 Percent Similarity: 98.41% Conservative: 0  
 Best Local Similarity: 98.41% Mismatches: 1  
 Query Match: 5.74% Indels: 2  
 DB: 12 Gaps: 0  
 US-09-787-097-12 (1-1429) x BI004303 (1-484)  
 QY 1183 ValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsn 1202  
 DB 75 GTGGCTACTCTGACGACAAACAGGAGTTTGACATGTTTCATCAATGCTCCCAAGATT 134  
 QY 1203 PheAsnLeuAsnIleThrTrrAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGlu 1222  
 DB 135 TTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAG 194  
 QY 1223 MetProVal-ValSerLysThrAsnIleLysGluTyrlYrLysAspSerPheSerAsnGlu 1242  
 DB 195 ATGCCTGC-TGTTTCAAAACCAACCATTAAGGAGTACAAAGATAGTTTCTTCTTAATGAGAA 253

```

QY 1242 sPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTr 1262
Db 254 GTTGTATTTCGACACCAACCAATATCACTTCTTTGTTATGTCAGTAATTCACCTG 313
QY 1262 pProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPh 1282
Db 314 GOCCATCAAAATTCAGATTGCCTTCTCTCAGCACAGCAATTTATGACCTGCTACAGTT 373
QY 1282 ePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTrpLy 1302
Db 374 CTTCGGACTTCTTCAGTGTGTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 433
QY 1302 sIleLysGlnSerCys 1307
Db 434 GATCAACAAGTGTGT 449

RESULT 56
CB520374
LOCUS CB520374 840 bp mRNA linear EST 09-JUL-2003
DEFINITION UI-W-GIO-csi-1-18-0-UI-r1 NIH_BMAP_GIO Mus musculus cDNA clone
IMAGE:6840067 5', mRNA sequence.
ACCESSION CB520374
VERSION CB520374.1 GI:29353729
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..840
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6840067"
/tissue_type="whole brain"
/dev_stages="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GIO"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with Ecor I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is ACCGAGACAG. This library was created for the University of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

FEATURES
source
ORIGIN
Alignment Scores:

```

```

Pred. No.: 5,33e-55 Length: 840
Score: 81.00 Matches: 81
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.67% Indels: 0
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CB520374 (1-840)

QY 842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861
Db 374 ATCAATCATCTCAAGATATGTCACAGCTCACTCTGACTCCATGGTTCGTTTCGAAG 433
QY 862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881
Db 434 ATCAATGTGTCTTACTGGTCTGGGAGGATATGTCCTCCATTCAAAATAGTTTGTGCAG 493
QY 882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901
Db 494 TGGATGCCATCTGAGCCAGTGAATGCTGCTTCTGTGGGATCTTGTGAGAGCTAGTACT 553
QY 902 ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
Db 554 CGGGATTAAAGCTGCAACCTGCACTCAACCTCTCAATGGCAGCGCTCTGTGAAGGCT 613
QY 922 Ala 922
Db 614 GCA 616

RESULT 57
R87660
LOCUS R87660 416 bp mRNA linear EST 16-AUG-1995
DEFINITION YP89G01.s1 Soares fetal liver spleen INFILS Homo sapiens cDNA clone
IMAGE:194640 3', mRNA sequence.
ACCESSION R87660
VERSION R87660.1 GI:946473
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 416)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and
Wilson, R.
The WashU-Merck EST Project
Unpublished (1995)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 1984
High quality sequence stops: 273
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert length: 1994 Std Error: 0.00
Seq primer: Promega -2lml3
High quality sequence stop: 273.
Location/Qualifiers
1..416
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3763690"
/db_xref="taxon:9606"
/clone="IMAGE:194640"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"

```

/clone lib="Soares fetal liver spleen INFLS"  
 /note="Organ: Liver and Spleen; Vector: pT7m3D (Pharmacia)  
 with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI;  
 1st strand cDNA was primed with a Pac I - oligo(dT) primer  
 15', AACGGAGAAATTAATAAGATCTTTTTTTTTTTTTTTT 3',  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Pac I and cloned into the Pac I  
 and Eco RI sites of the modified pT7m3 vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M.Patima Bonaldo."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,74e-54 Length: 416  
 Score: 80.00 Matches: 80  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.60% Indels: 0  
 DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x R87660 (1-416)

Qy 753 MetTyrCysAsnLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGln 772  
 Db 1 ATGTACTACTGTAAACAGAACACAGCTGCAGAGCTGCGCTCGACACAGATGCCAG 60  
 Qy 773 TTPGIUProArgAsnGlnCysileAlaLeuProGluAsnileCysGlyleGlyTTP 792  
 Db 61 TGGGAGCCCCGGAATCAGAGGTGATGCCCTGCCGGAATAATCTGTGGCATTCGCTGG 120  
 Qy 793 HisLeuValGlyAsnSerCysLeuLysIleThrAlaLysGluAsnTyrAspAsnAla 812  
 Db 121 CATTTGGTTGGAAATCATGTTTGAATAATCTACTGCGCAAGAGAGATATGACAATGCT 180  
 Qy 813 LysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrGlnLysLysVal 832  
 Db 181 AAATGTCTGTAGAACACCAATGCCCTTTGGCTTCTCTACACCCAGAGAGGTA 240

## RESULT 58

BF854487/c  
 LOCUS BF854487 567 bp mRNA linear EST 16-JAN-2001  
 DEFINITION MR2-EN0093-271200-005-a04 EN0093 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF854487  
 VERSION BF854487.1 GI:124242231  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE  
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,  
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

PUBMED

COMMENT

Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?ci=MR2&t2=MR2-EN0093-  
 271200-005-a04&t3=2000-12-27&t4=1)

Seq primer: puc 18 forward  
 High quality sequence stop: 566.

## FEATURES

source Location/Qualifiers  
 1..567  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev stage="Adult"  
 /clone lib="EN0093"  
 /note="Organ: lung normal; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,35e-54 Length: 567  
 Score: 80.00 Matches: 166  
 Percent Similarity: 97.65% Conservative: 0  
 Best Local Similarity: 97.65% Mismatches: 2  
 Query Match: 5.60% Indels: 4  
 DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BF854487 (1-567)

Qy 1096 PheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAla 1115  
 Db 567 TTCTACGGTGATCCACCAATGAGGGAATGTGACCCATGCAAGTGCATGGCAGCG 508  
 Qy 1116 SerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAsp 1135  
 Db 507 TCTCTGTGCAACCAACACGGGCAAGTGCTTCTGCACCACCAAGGGCGTCAAGGGGGAC 448  
 Qy 1136 GlnCysGlnLeuCysGluValGluAsnArgTyrGlnGly-AsnProLeuArgGlyThrCy 1155  
 Db 447 GAGTGCCAGCTATGTGAGGTAGAAATCGATACCAAGA-AAACCCCTCTCAGAGGAACATG 389  
 Qy 1155 sTyrTyrThrLeuLeuileAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspAr 1175  
 Db 388 TTATTACTCTCTTATTGACTATCAGTTACCTTTAGTCTATCCCGAGGAGATGATCG 329  
 Qy 1175 gTyrTyrThrAlaIleAsnPheValIaThrProAspGluGlnAsnArgAspLeuAspMe 1195  
 Db 328 CTATTACACAGCTATCAATTTTGTGGCTACTCTCGACGAACAAACAGGATTTGGACAT 269  
 Qy 1195 tPheileAsnAlaSerLysAsnPheAsnLeuAsnIleThrTyrAlaAlaSerPheSerAl 1215  
 Db 268 GTTCATCAATGCCTCCAGAAATTCACCTCAACATCACCCTGGCTGCCAGTTCTCAGC 209  
 Qy 1215 aGlyThrGlnAlaGlyGluGluMetProValValSer-LysThrAsnIleLysGluTyrL 1235  
 Db 208 TGGAAACCAGGCTGGAGAGAGATCCCTGTGTTCAAAAGCC-AACATTAAAGAGATACA 150  
 Qy 1235 ysAspSerPheSerAsnGluLysPheAspPheArgenHisProAsnIleThrPhePheV 1255  
 Db 149 AAGATAGTTTCTCTAATGAGAAGTTGATTTTCGACACCAACCAATATCATTCTTTTG 90  
 Qy 1255 alTyrValSerAsnPheThrTyrPro 1263  
 Db 89 TTTATGTACAGTAATTTTCCCTGGCCC 64

## RESULT 59

AW579586  
 LOCUS AW579586 427 bp mRNA linear EST 16-MAR-2000  
 DEFINITION MRO-HT0209-100100-103-e09 HT0209 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION AW579586  
 VERSION AW579586.1 GI:7254635  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 427)  
 HCGP <http://www.ludwig.org.br/ORESTES>.  
 The FAPESP/LICR Human Cancer Genome Project  
 Unpublished (1999)  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR0&t2=MR0-HT0209-100100-103-e09&t3=2000-01-10&t4=1>)  
 Seq primer: puc 18 forward  
 High quality sequence stop: 427.

#### FEATURES

source  
 1. .427  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="HT0209"  
 /note="Organ: head neck; Vector: puc18; Site\_1: SmaI;  
 Site\_2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the pUC 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

#### ORIGIN

Alignment Scores:  
 Pred. No.: 1,156-53 Length: 427  
 Score: 79.00 Matches: 115  
 Percent Similarity: 98.29% Conservative: 0  
 Best Local Similarity: 98.29% Mismatches: 1  
 Query Match: 5.53% Indels: 2  
 DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x AW579586 (1-427)

QY 434 LeuThrProValAlaValGlnTyrAlaValValGlyHisSerAlaHisLeuValThr 453  
 DB 64 TTGACCCCTTAAGCAAGGAGGAGTATGCGTGGTGGGACCTCTGCACATTTGTACA 123  
 QY 454 LeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyr 473  
 DB 124 CTGAGAATGGCGAGTGGTCTATCTGGTCACTCTTGGTCACTGCCCTCTCTATGGATAT 183  
 QY 474 IleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGln 493  
 DB 184 ATAAGCAATGTGCAGAAATATATTGGATAAGACACATGGAGTATTATACACACCCAG 243  
 QY 494 GlyAlaLeuValGlnGlyGlyTyrGlyHisSerValTyrAspHisArgThrArgala 513  
 DB 244 GGTGCCCTTGTGCARAGGGGGTTACGCCCATAGCAGTGTTCAGACCATAGGACCGGG- 302  
 QY 514 -LeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAs 533  
 DB 303 CCTATACGTTTATGTTGGTGTACAGGCTTTTACAGTCCCAATAGTACCGGCTTGCAGATGA 362  
 QY 533 pLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAsp 549  
 DB 363 TCTTACCGATATGATGTGGATACCCAGATGTGGACCATTTCTTAAGGAC 411

RESULT 60  
 BI202540  
 LOCUS

BI202540 434 bp mRNA linear EST 29-APR-2002

DEFINITION UI-M-CGOp-brh-h-01-0-UI.s1 NIH\_BMAP\_Ret4\_S2 Mus musculus cDNA clone  
 UI-M-CGOp-brh-h-01-0-UI 3', mRNA sequence.  
 ACCESSION BI202540  
 VERSION B1202540.1 GI:14668512  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 434)  
 Bernaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genom Res. 6 (9), 791-806 (1996)  
 JOURNAL 97044477  
 MEDLINE 8889548  
 PUBMED  
 COMMENT  
 Contact: Chin, H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: [mbest@mail.nih.gov](mailto:mbest@mail.nih.gov)  
 Oligo-dT track not found, Not 1 site shown in beginning of sequence  
 is likely internal to the message. cDNA Library Preparation: M.B.  
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA  
 clones from RESEARCH GENETICS. It should be noted that Bento Soares  
 is generating a small number of additional specialized  
 non-redundant arrays of BMAP cDNAs whose availability will be  
 considered under appropriate and limited collaborative arrangements  
 The tissue for this library was contributed by Dr. Xin-Yuan Fu,  
 Yale University School of Medicine  
 Seq primer: M13 Forward  
 POLYA=No.

#### FEATURES

Location/Qualifiers  
 1. .434

/organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-CGOp-brh-h-01-0-UI"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NIH\_BMAP\_Ret4\_S2"  
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; The  
 NIH BMAP Ret4 S2 library is a subtracted library,  
 ultimately derived from mouse retina tissue libraries at  
 various stages of development. For a detailed description  
 of the library from which this clone was derived, please  
 visit our web site at [brainest.eng.uiowa.edu](http://brainest.eng.uiowa.edu). The tissue  
 for this library was contributed by Dr. Xin-Yuan Fu, Yale  
 University School of Medicine  
 TAG\_SEQ=None found"

#### ORIGIN

Alignment Scores:  
 Pred. No.: 3,166-51 Length: 434  
 Score: 76.00 Matches: 76  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.32% Indels: 0  
 DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BI202540 (1-434)

QY 123 GlyTrpValGlyGluGlnCysGlnHisCysGlyArgPheArgLeuThrGlySerSer 142  
 DB 45 GGCTGGTGGCGAGCAATGCCAGCACTGCCGGGGCGCTTCAGACTAACTGGCTCTCT 104  
 QY 143 GlyPheValThrAspGlyProGlyAsnTyrLysTyrThrLysCysThrTrpLeuIle 162  
 DB 105 GGATTTGTAAACATGACCTGGGAATTATAATATAGCAAGTGCATGGCTCATTT 164

Qy 163 GluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSer 182  
 Db 165 GAAGACAGCCAAATAGAAATAGAGCTTCCTTCAACCTTTTGTCTACAGATGAGC 224  
 Qy 183 TrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198  
 Db 225 TGGGACCAATTATATGTTATGATGGGACTCAATCTACGCCTCTG 272

RESULT 61  
 AW120845  
 LOCUS  
 DEFINITION  
 UI-M-BH2.3-aca-e-06-0-UI.s1 NIH BMAP M.S3.3 Mus musculus cDNA clone  
 UI-M-BH2.3-aca-e-06-0-UI.3', mRNA sequence.  
 ACCESSION  
 AW120845  
 VERSION  
 AW120845.1 GI:6096178  
 KEYWORDS  
 EST.  
 SOURCE  
 Mus musculus (house mouse)  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 TITLE  
 Genome Res. 6 (9), 791-806 (1996)

97044477  
 889548  
 Contact: Chin, H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA

Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: mEST@mail.nih.gov  
 Oligo-qr track not found, Not I site shown in beginning of sequence  
 is likely internal to the message. cDNA Library Preparation: M.B.  
 Soares Lab Clone distribution: NIH BMAP cDNA clones will be made  
 available by the means that is soon to be determined. When NIH  
 determines the means for distribution of the BMAP cDNA clones, this  
 record will be updated accordingly when that means is determined.  
 Seq primer: M13 Forward  
 POLYA=No.

# FEATURES

source

Location/Qualifiers  
 1..435  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="CS7BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-BH2.3-aca-e-06-0-UI"  
 /dev\_stage="27-32 days"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NIH BMAP M.S3.3"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site: Not I; Site 2: Eco RI; The  
 NIH BMAP M.S3.3 library is a subtracted library of a  
 series, ultimately derived from a mixture of individually  
 tagged, normalized libraries from ten regions of the mouse  
 brain (cerebellum, brain stems, olfactory bulbs,  
 hypothalamus, cortex, amygdala, basal ganglia, pineal  
 gland, striatum, hippocampus) after a series of  
 subtractions to reduce the representation of cDNAs from  
 which ESTs had already been generated. The following  
 serially subtracted libraries were generated in this  
 process: NIH BMAP M.S3.3, NIH BMAP M.S2, NIH BMAP M.S1.  
 The subtracted library (NIH BMAP M.S3.3) was constructed  
 as follows: PCR amplified cDNA inserts from NIH BMAP M.S2  
 clones from which 3' ESTs had been derived was used as a  
 driver in a hybridization with the NIH BMAP M.S2 library  
 in the form of single-stranded circles. The remaining  
 single-stranded circles (subtracted library) was purified  
 by hydroxyapatite column chromatography, converted to  
 double-stranded circles and electroporated into DH10B  
 bacteria (Life technologies) to generate the

## ORIGIN

Alignment Scores: 3.17e-51 Length: 435  
 Pred. No.: 76.00 Matches: 76  
 Score: 100.00% Conservative: 0  
 Percent Similarity: 100.00% Mismatches: 0  
 Best Local Similarity: 5.32% Indels: 0  
 Query Match: 10 Gaps: 0  
 DB:

US-09-787-097-12 (1-1429) x AW120845 (1-435)

Qy 123 GLYTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSer 142  
 Db 44 GCGTGGGTGGCGGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAACTGGCTCTCT 103  
 Qy 143 GlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeu 162  
 Db 104 GGATTTGTAAACAGATGGACCTGGGAATTATAAATATAAGACGAAGTGCACATGGCTCA 163  
 Qy 163 GluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSer 182  
 Db 164 GAAGACAGCCAAATAGATAGACTTCGCTTCAACATTTTGTCTACAGATGAGC 223  
 Qy 183 TrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198  
 Db 224 TGGGACCAATTATATGTTATGATGGGACTCAATCTACGCCTCTG 271

## RESULT 62

BI202684

LOCUS

DEFINITION

UI-M-CG0p-brk-h-05-0-UI.s1 NIH BMAP Rat4.S2 Mus musculus cDNA clone

UI-M-CG0p-brk-h-05-0-UI.3', mRNA sequence.

ACCESSION

BI202684

VERSION

BI202684.1 GI:14668656

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 435)

AUTHORS

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

JOURNAL

Genome Res. 6 (9), 791-806 (1996)

MEDLINE

97044477

PUBMED

889548

COMMENT

Contact: Chin, H

National Institute of Mental Health

6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD

20892-9643, USA

Tel: 301 443 1706

Fax: 301 443 9890

Email: mEST@mail.nih.gov

Oligo-qr track not found, Not I site shown in beginning of sequence

is likely internal to the message. cDNA Library Preparation: M.B.

Soares Lab Clone distribution: Researchers may obtain BMAP cDNA

clones from RESEARCH GENETICS. It should be noted that Benco Soares

is generating a small number of additional specialized

non-redundant arrays of BMAP cDNAs whose availability will be

considered under appropriate and limited collaborative arrangements

The tissue for this library was contributed by Dr. Xin-Yuan Fu,

Yale University School of Medicine

Seq primer: M13 Forward

POLYA=No.

LOCATION/Qualifiers

1..435





LOCUS BE952846 566 bp mRNA linear EST 04-OCT-2000  
 DEFINITION UI-M-BH4-azk-g-05-0-UI.s1 NIH BMAP\_M S5 Mus musculus cDNA clone  
 UI-M-BH4-azk-g-05-0-UI 3', mRNA sequence.  
 ACCESSION BE952846  
 VERSION BE952846.1 GI:10593734  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 566)  
 AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 889548  
 COMMENT Contact: Chin, H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: mEST@mail.nih.gov  
 Oligo-dT track not found, Not 1 site shown in beginning of sequence  
 is likely internal to the message. cDNA Library Preparation: M.B.  
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA  
 clones from RESEARCH GENERICS. It should be noted that Bento Soares  
 is generating a small number of additional specialized  
 non-redundant arrays of BMAP cDNAs whose availability will be  
 considered under appropriate and limited collaborative arrangements  
 Seq primer: M13 Forward  
 POLYA=No.

## FEATURES

source

Location/Qualifiers

1..566  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-BH4-azk-g-05-0-UI"  
 /dev\_stage="27-32 days"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NIH BMAP M S5"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not 1; Site 2: Eco RI; The  
 NIH\_BMAP\_M\_S5 library is a subtracted library of a series,  
 ultimately derived from a mixture of individually tagged,  
 normalized libraries from ten regions of the mouse brain  
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,  
 cortex, amygdala, basal ganglia, pineal gland, striatum,  
 hippocampus) after a series of subtractions to reduce the  
 representation of cDNAs from which ESTs had already been  
 generated. For a detailed description of the library from  
 which this clone was derived, please visit our web site  
 at brainest.eng.uiowa.edu.  
 TAG\_SEQ=None found"

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,1e-51 Length: 566  
 Score: 76.00 Matches: 76  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5,32% Indels: 0  
 DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BE952846 (1-566)

QY 123 GlyTTPValGlyGluGlnCysGlnHisCysGlyArgPheArgLeuThrGlySerSer 142  
 Db 44 GCCTGGTGGGGAGCAATGCCAGCACTGCGGGGGCGCTTCAGACTAACTGGCTCTTCT 103  
 QY 143 GlyPheValThrAspGlyProGlyAsnTyrlsTyrlsLysCysThrTrpLeuIle 162

Db 104 GCATTGTACAGATGACCTGGAAATTATAAATAGCAAGTGCACATGGCTCAT 163  
 QY 163 GluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSer 182  
 Db 164 GAAGGACAGCCCAATAGAAATATGACATTCGCTTCACCACTTGTCTACAGATGAC 223  
 QY 183 TTPAspHisLeuTyrlsValTyrlsAspGlyAspSerIleTyrlsAlaProLeu 198  
 Db 224 TGGGACCAATTATATGTTTATGATGGGACTCAATCTAGCACCTCTG 271  
 RESULT 65  
 BE980188  
 LOCUS  
 DEFINITION UI-M-BG2-bcq-f-12-0-UI.s1 NIH BMAP MSC S1 Mus musculus cDNA clone  
 UI-M-BG2-bcq-f-12-0-UI 3', mRNA sequence.  
 ACCESSION BE980188  
 VERSION BE980188.1 GI:10647894  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 626)  
 AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 889548  
 COMMENT Contact: Chin, H  
 National Institute of Mental Health  
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
 20892-9643, USA  
 Tel: 301 443 1706  
 Fax: 301 443 9890  
 Email: mEST@mail.nih.gov  
 Oligo-dT track not found, Not 1 site shown in beginning of sequence  
 is likely internal to the message. cDNA Library Preparation: M.B.  
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA  
 clones from RESEARCH GENERICS. It should be noted that Bento Soares  
 is generating a small number of additional specialized  
 non-redundant arrays of BMAP cDNAs whose availability will be  
 considered under appropriate and limited collaborative arrangements  
 Seq primer: M13 Forward  
 POLYA=No.

## FEATURES

source

Location/Qualifiers

1..626  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="UI-M-BG2-bcq-f-12-0-UI"  
 /dev\_stage="27-32 days"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NIH BMAP MSC S1"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not 1; Site 2: Eco RI; The  
 NIH\_BMAP\_MSC\_S1 library is a subtracted library derived  
 from which this clone was derived, please visit our web  
 site at brainest.eng.uiowa.edu.  
 TAG\_SEQ=None found"

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,52e-51 Length: 626  
 Score: 76.00 Matches: 76  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5,32% Indels: 0  
 DB: 10 Gaps: 0

```

US-09-787-097-12 (1-1429) x BE980188 (1-626)

QY 123 GlyTTPValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSer 142
Db 44 GGCTGGGTGGGCGAGCATGCCAGCACTCGCGGGGCCCTTCAGACTAACTGGCTCTCT 103
QY 143 GlyPheValThrAspGlyProGlyAsnTyrLysTyrLysCysThrTrpLeuLe 162
Db 104 GGATTGTGAACAGATGGACCTGGGAATTATTAATATAAGACCAAGTGCACATGGCTCAT 163
QY 163 GluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSer 182
Db 164 GAAGGACACGCCAAATAGATAATAGACTTCGCTTCAACCACTTTTGCTACAGATGTAGC 223
QY 183 TrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198
Db 224 TGGGACCACTTATATGTTTATGATGGGACTCACTACGACCTCTG 271

RESULT 66
BE671400/c
LOCUS
DEFINITION
750d11.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:3285909 3'
similar to TR_060295 O60295 KIAA0548 PROTEIN ; contains L1.b1 L1
repetitive element 1, mRNA sequence.
BE671400
445 bp mRNA linear EST 08-SEP-2000
BE671400.1 GI:10031941
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 445)
REFERENCE
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 397.
FEATURES
Location/Qualifiers
1..445
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3285909"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu24"
/note="Organ: lung; Vector: pTVT3D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI-CGAP Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneIDs
1414920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."
ORIGIN
Alignment Scores: 2.1e-50 Length: 445
Pred. No.: 75.00 Matches: 75
Score: 100.00% Conservative: 0
Percent Similarity: 100.00%

```

```

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.25% Indels: 0
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BE671400 (1-445)

QY 1193 LeuAspMetPheIleAsnAlaSerIysAsnPheAsnLeuAsnIleThrTrpAlaLaser 1212
Db 445 TTGGCATGTTTCATCAATGCCTCCAGAAATTTCAACCTCAACATCACCTGGCTGCCAGT 386
QY 1213 PheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerIysThrAsnIleLys 1232
Db 385 TTCTCAGCTTGACACCGCTGGAGAGAGATGCCGTGTTGTTTCARAAACCAACATTAG 326
QY 1233 GluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThr 1252
Db 325 GAGTACAAAGATAGTTTCTCTAATGAGAAATTTGATTTTCGCAACCAACCAATCACT 266
QY 1253 PhePheValTyrValSerAsnPheThrTrpProIleLysIleGln 1267
Db 265 TTCTTTGTTTATGTCAGTAATTTACCTGCGCCATCAAAATTCAG 221

RESULT 67
AI076479/c
LOCUS
DEFINITION
OZ28B04.x1 Soares total fetus Nb2HF8.9w Homo sapiens cDNA clone
IMAGE:167623 3' similar to contains L1.b1 L1 repetitive element 1,
mRNA sequence.
AI076479
464 bp mRNA linear EST 01-OCT-1998
AI076479.1 GI:3405657
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 464)
REFERENCE
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1450 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 444.
FEATURES
Location/Qualifiers
1..464
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:167623"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_lib="Soares total fetus Nb2HF8.9w"
/note="Vector: pTVT3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not 1; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5',
TGTTACCATCTGAGTGGAGCGGCGCTTAATTTTATTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pTVT3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
ORIGIN
Alignment Scores: 1.41e-49 Length: 464
Pred. No.: 74.00 Matches: 74
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00%

```

Query Match: 5.18% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AI076479 (1-464)

QY 1194 AspMetPheIleAsnAlaSerLysAsnLysPheAsnLeuAsnIleThrTTPAlaAlaSerPhe 1213  
DB 463 GACATGTTTCATCAATGCTCCAGAAATTTCACTCACAATCAGTGGCTGCCAGTTTC 404

QY 1214 SerAlaGlyThrGlnAlaGlyGluMetProValSerLysThrAsnIleLysGlu 1233  
DB 403 TCAGTGGAAACCCAGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAAGAG 344

QY 1234 TyrIysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhe 1253  
DB 343 TACAAGATAGTTTCTTAATGAGAGATTTGATTTTGCACACCCCAAAATATCACTTTC 284

QY 1254 PheValTyrValSerAsnPheThrTTPProIleLysIleGln 1267  
DB 283 TTTGTTTATGTCAGTAATTTCCCTGGCCCATCAAAATTCAG 242

RESULT 68  
BF549645  
LOCUS  
DEFINITION BF549645 474 bp mRNA linear EST 12-DEC-2000  
UI-R-C2-nf-e-05-0-UI.r1 UI-R-C2 Rattus norvegicus cDNA clone  
UI-R-C2-nf-e-05-0-UI 5', mRNA sequence.

ACCESSION BF549645.1 GI:11659375  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Rattus norvegicus (Norway rat)  
Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus

REFERENCE  
1 (bases 1 to 474)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)

JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@iowa.edu  
cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
clones will be available through Research Genetics (www.resgen.com)  
This clone is also available through the I.M.A.G.E. Consortium at  
LLNL (info@image.llnl.gov). IMAGE ID= 1776329 The following  
repetitive elements were found in this cDNA sequence: 439-465,  
>POLY A#Simple repeat  
Seq primer: M13 Forward.  
Location/Qualifiers  
1..474  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-C2-nf-e-05-0-UI"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker. Site 1: Not 1; Site 2: Eco RI; The UI-R-C2  
library is a subtracted library derived from the UI-R-C1  
library, which is a subtracted library derived from the  
UI-R-C0 library. The UI-R-C0 library consisted of a  
mixture of individually tagged normalized libraries  
constructed from rat placenta, adult lung, brain, liver,  
kidney, heart, spleen, ovary, muscle, 8, 12 and 18-day

embryo. The tag is a string of 3-5 nucleotides present  
between the Not I site and the oligo-dT track which allows  
identification of the library of origin of a clone within  
the mixture. The subtracted library (UI-R-C2) was  
constructed as follows: PCR amplified cDNA inserts from  
UI-R-C1 clones from which 3' ESTs had been derived was  
used as a driver in a hybridization with the UI-R-C1  
library in the form of single-stranded circles. The  
remaining single-stranded circles (subtracted library) was  
purified by hydroxyapatite column chromatography,  
converted to double-stranded circles and electroporated  
into DH10B bacteria (Life Technologies) to generate the  
UI-R-C2 library. This procedure has been previously  
described (Bonaldo, Lennon and Soares, Genome Research 6:  
791-806, 1996)"

ORIGIN

Alignment Scores: 9.33e-49 Length: 474  
Pred. No.: 73.00 Matches: 73  
Score: 73.00 Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 5.11% Gaps: 0  
DB: 10

US-09-787-097-12 (1-1429) x BF549645 (1-474)

QY 849 SerLysLeuThrLeuThrProTTPValGlyLeuAlaGlyLeuValSerTTPCys 868  
DB 20 TCCAAGTCACTCTGACTCATGGTGGTTCGGAAGATCAATGATCTTACTGTGTC 79

QY 869 TTPGluAspMetSerProPheThrAsnSerLeuLeuGlnTTPMetProSerGluProSer 888  
DB 80 TGGAGGATATGTCCTCCATTCAAAATAGTTTGTCTGATGGATGCTATCTGAGCCAGT 139

QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908  
DB 140 GATGCGCGGCTCTGTGGATCTTGTGAGAGCTTAGTACTCGGGGACTGAAGGCTGCAACC 199

QY 909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
DB 200 TGCATCAACCCACTCAATGGAGTGTGTGAAAGGCGCT 238

RESULT 69  
BE865031  
LOCUS  
DEFINITION BE865031 535 bp mRNA linear EST 29-SEP-2000  
UI-M-BH1-amt-f-11-0-UI.r1 NIH BMAP M.92 Mus musculus cDNA clone  
UI-M-BH1-amt-f-11-0-UI 5', mRNA sequence.

ACCESSION BE865031.1 GI:10386668  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 535)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)

JOURNAL  
MEDLINE  
PUBMED  
COMMENT  
Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: hest@mail.nih.gov  
cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It  
should be noted that Bento Soares is generating a small number of  
additional specialized non-redundant arrays of BMAP cDNAs whose

availability will be considered under appropriate and limited collaborative arrangements

# FEATURES

Seq primer: M3 Reverse.  
Location/Qualifiers  
1. .535

/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UI-M-BH1-anc-f-11-0-UI"  
/dev\_stage="27-32 days"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="NIH BMAP M S2"  
/note="Vector: pTT3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; The NIH BMAP M S2 library is a subtracted library derived from a mixture of normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus). The driver used for subtraction consisted of a pool of 5,000 clones from the NIH BMAP M S1 library and a pool of 2,000 clones obtained from non-normalized and normalized mouse brain spinal cord libraries."

## ORIGIN

Alignment Scores:  
Pred. No.: 1,848-45 Length: 535  
Score: 69.00 Matches: 125  
Percent Similarity: 98.43% Conservativeness: 0  
Best Local Similarity: 98.43% Mismatches: 1  
Query Match: 4.83% Indels: 2  
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x BE865031 (1-535)

QY 1083 LeuThrThrGlyLysHisCysGluThrCysLeuSerGlyPheTyrGlyAspProThrAsn 1102  
DB 1 CTGACCACGGGCAAGCACTGCAGACCTGCATATCTGGCTTCTATGGTGGACCGCACTAAT 60  
QY 1103 GlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThr 1122  
DB 61 GGAGGCAATGTCACCATCGAAGTCAGATGGGACCATCTACTGTGCACACCAACACC 120  
QY 1123 GlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluVal 1142  
DB 121 GGCAGTGTCTTCTGTACACCAAGGTGTCAAGGGGACGAGTGCAGCTATGTGAGGTA 180  
QY 1143 GluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrLeuLeuLeuAs 1162  
DB 181 GAAATCGATACACGGGAACCTCTCAA-AGGAACATGCTACTATACCTCTCATTTGA 239  
QY 1162 pTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIleAsnPh 1182  
DB 240 CTATCAGTTCACTTTAGCTGTCTCCAGGAAGACCGCTACTACACAGCCATCACTT 299  
QY 1182 eValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAs 1202  
DB 300 TGTGCTACTCTGTATGACAAACAGGATTTGACATGTTCAATGCTCCCAAAA 359  
QY 1202 nPheAsnLeuAsnIleThr 1208  
DB 360 CTTCACTTCACATCACC 378

RESULT 70  
A1818750/c  
LOCUS  
DEFINITION  
w111c03.x1 NCI CGAP Ut1 Homo sapiens cDNA clone IMAGE:2424580 3', similar to TR:060295 O60295 KIAA0548 PROTEIN ;contains L1.L3 L1 repetitive element ;, mRNA sequence.  
ACCESSION  
A1818750  
VERSION  
A1818750.1 GI:5437829

## KEYWORDS

Homo sapiens (human)  
Homo sapiens

## SOURCE

ORGANISM

## REFERENCE

1 (bases 1 to 589)

## AUTHORS

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

## TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

## JOURNAL

Unpublished (1997)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgapsb@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html  
Insert Length: 2056 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 399.  
Location/Qualifiers  
1. .589

## FEATURES

source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2424580"  
/tissue\_type="well-differentiated endometrial adenocarcinoma, 7 pooled tumors"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Ut1"  
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.75 kb. Life Technologies catalog #: 11538-014"

## ORIGIN

Alignment Scores:  
Pred. No.: 2,028-45 Length: 589  
Score: 69.00 Matches: 112  
Percent Similarity: 98.25% Conservativeness: 0  
Best Local Similarity: 98.25% Mismatches: 1  
Query Match: 4.83% Indels: 2  
DB: 9 Gaps: 0  
US-09-787-097-12 (1-1429) x A1818750 (1-589)  
QY 1155 CysTyrTyrThrLeuLeuLeuLeuLeuPheThrPheSerLeuSerGlnGluAspAsp 1174  
DB 556 TGTATTATATCTCTTCTTATGACTATATGATTCACCTTCAGTCTATCCACGAGATGAT 497  
QY 1175 ArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgPheLeuAsp 1194  
DB 496 CGCTATTACACAGTATCAATTTTGTGGCTACTCTCTGACGAAACAAACAGGGATTTGGAC 437  
QY 1195 MetPheIleAsn-AlaSerLysAsnPheAsnLeuAsnIleThrTyrAlaAlaSerPheSe 1214  
DB 436 ATGTTTCATCCA-TGCTCTCAAGAAATTTCAACCTCAACATCACCCTGGCGTCCGAGTTTCTC 378  
QY 1214 rAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyr 1234  
DB 377 AGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAAGAGTA 318  
QY 1234 rLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePh 1254  
DB 317 CAAAGATAGTTTCTCTAATGAGAGATTTGATTTTCGACACCAACCAATATATCATTCTT 258  
QY 1254 eValTyrValSerAsnPheThrTyrProIleLysIleGln 1267  
DB 257 TGTATTATGTCAGTAATTTTCACTGGCCCAATCAAAATTCAG 218

```

RESULT 71
BB650220/c
DEFINITION
  UI-M-BH2.1-apr-b-08-0-UI.r1 NIH_BMAP_M.S3.1 Mus musculus cDNA clone
  UI-M-BH2.1-apr-b-08-0-UI 5', mRNA sequence.
ACCESSION
  BE650220
VERSION
  BE650220.1 GI:9976044
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 714)
AUTHORS
  Ronaldo, M.F., Lennon, G. and Soares, M.B.
TITLE
  Normalization and subtraction: two approaches to facilitate gene
  discovery
JOURNAL
  Genome Res. 6 (9), 791-806 (1996)
MEDLINE
  97044477
PUBMED
  8889548
COMMENT
  Contact: Chin, H
  National Institute of Mental Health
  6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
  20892-9643, USA
  Tel: 301 443 1706
  Fax: 301 443 9890
  Email: m8est@mail.nih.gov
  cDNA Library Preparation: M.B. Soares Lab Clone distribution:
  Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It
  should be noted that Bento Soares is generating a small number of
  additional specialized non-redundant arrays of BMAP cDNAs whose
  availability will be considered under appropriate and limited
  collaborative arrangements
  Seq primer: M13 Reverse.
  Location/Qualifiers
    1..714
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="CS7BL/6J"
    /db_xref="taxon:10090"
    /clone="UI-M-BH2.1-apr-b-08-0-UI"
    /dev_stage="27-32 days"
    /lab_host="DH10B (Life Technologies)"
    /clone_lib="NIH_BMAP_M.S3.1"
    /note="Vector: pT73D-Pac (Pharmacia) with a modified
    polylinker; Site 1: Not I; Site 2: Eco RI; The
    NIH_BMAP_M.S3.1 library is a subtracted library of a
    series, ultimately derived from a mixture of individually
    tagged normalized libraries from ten regions of the mouse
    brain (cerebellum, brain stems, olfactory bulbs, pineal
    hypothalamus, cortex, amygdala, basal ganglia, pineal
    gland, striatum, hippocampus) after a series of
    subtractions to reduce the representation of cDNAs from
    which ESTs had already been generated. The following
    serially subtracted libraries were generated in this
    process: NIH_BMAP_M.S3.1, NIH_BMAP_M.S2, NIH_BMAP_M.S1.
    The subtracted library (NIH_BMAP_M.S3.1) was constructed
    as follows: PCR amplified cDNA inserts from NIH_BMAP_M.S2
    clones from which 3' ESTs had been derived was used as a
    driver in a hybridization with the NIH_BMAP_M.S2 library
    in the form of single-stranded circles. The remaining
    single-stranded circles (subtracted library) was purified
    by hydroxyapatite column chromatography, converted to
    double-stranded circles and electroporated into DH10B
    bacteria (Life Technologies) to generate the
    NIH_BMAP_M.S3.1 library. This procedure has been
    previously described (Bonaldo, Lennon and Soares, Genome
    Research 6:791-806, 1996)"
  ORIGIN
    Alignment Scores:
    Pred. No.: 2.43e-45 Length: 714
    Score: 69.00 Matches: 172
    Percent Similarity: 97.73% Conservative: 0

Best Local Similarity: 97.73% Mismatches: 2
Query Match: 4.83% Indels: 4
DB: 10 Gaps: 0
US-09-787-097-12 (1-1429) x BE650220 (1-714)
QY 1037 TTTPrGlnPrLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056
Db 585 TATCCAGAGCCCTTCTGAACTCCAGATGTTGTAGAGACAGCAGATACAACTGGTCT 526
QY 1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076
Db 525 TTCAATCACTGTCCAGTTGCCAGTCCAAACGACACAGCAATGCAATCAACAGAGTATC 466
QY 1077 CysGluLysCysGluAsn-LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPh 1096
Db 465 TGTGAGAAGTGTGAGGA-CCTGACACAGGGCAGCACTCGGAGACCTGCATATCTGGCTT 407
QY 1096 eTyrGlyAspProThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSe 1116
Db 406 CTATGTTGACCCGACTAATGGAGGCAATGTGACCCATGCAATGCAATGGCAGCATC 347
QY 1116 rLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspG1 1136
Db 346 ACTGTGCACACCAACACCGCGCAAGTCTTCTGTACCAACAAAGGTGTCAAGGGGACGA 287
QY 1136 uCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysT 1156
Db 286 GTCCAGCATGTGTAGGTAGAAATCGATACAGGCAAAACCTCTCTCAA-AGGAACATGCT 228
QY 1156 YrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgT 1176
Db 227 ACTATACCTTCTCATTTGACTATCATCTTACCTTTAGCTGTCCAGGAGAGACGCGCT 168
QY 1176 YrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetP 1196
Db 167 ACTACACAGCCATCAACTTTGTGGCTACTCTCTGATGAACAAACAGGGAATTGACATGT 108
QY 1196 heIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAla 1210
Db 107 TCATCAATGCTCCAAAACTTCAAGCTCAACATCACTGGGCC 64

RESULT 72
CD804170
LOCUS
  UI-M-GV0-cie-b-08-0-UI.r1 NIH_BMAP_GV0 Mus musculus cDNA clone
  IMAGE:30545215 5', mRNA sequence.
ACCESSION
  CD804170
VERSION
  CD804170.1 GI:32462996
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
REFERENCE
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 292)
AUTHORS
  NIH-WGC http://wgc.nci.nih.gov/
TITLE
  National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
  Unpublished (1999)
COMMENT
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-x@mail.nih.gov
  Tissue Procurement: Dr. Jim Lin, University of Iowa
  cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
  DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
  Clone Distribution: Distribution information can be found at
  http://genome.uiowa.edu/distribution/mousefl.html
  This clone was contributed by the Brain Molecular Anatomy Project
  (BMAP)
  Seq primer: pyX-5.
  Location/Qualifiers
    1..292
    /organism="Mus musculus"
    /mol_type="mRNA"

FEATURES
  source

```

/strains="C57BL/6"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:30545215"  
 /tissue\_type="whole brain"  
 /dev\_stage="1,5, and 15 days newborn"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_lib="NIH BMAP GV0"  
 /note="Organ: Brain; Vector: pYX-Asc; Site\_1: EcoR I; Site\_2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAAGCTGAAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., Program coordinator."

## ORIGIN

Alignment Scores:  
 Pred. No.: 6,588-45 Length: 529  
 Score: 68.00 Matches: 292  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.76% Indels: 0  
 DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CD804170 (1-292)

QY 842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTropValGlyLeuArgLys 861  
 Db 87 ATGCAATCATCTCAAGTAATGTCAGGCTCATCTGACTTCAAGGTTGGTCTTCGGAAG 146  
 QY 862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881  
 Db 147 ATCAATGTCTTACTGTGGTGGGAGGATATGCTCCATTACAAATAGTTTGTGCGAG 206  
 QY 882 TrpMetProSerGlnProSerAspLaglyPheCysGlyIleLeuSerGlnProSerThr 901  
 Db 207 TGGATGCCATCTGAGCCCATGATGCTGCTTCTGTGGGCTTGTTCAGAGCCCTAGTACT 266  
 QY 902 A-GGlyLeuLysAlaAlaThrCys 909  
 Db 267 CGGGGATTAAAGGCTGCAACCTGC 290

RESULT 73  
 BM537406  
 LOCUS  
 DEFINITION 529 bp mRNA linear EST 20-FEB-2002  
 clone ha82a01 5', mRNA sequence.

ACCESSION BM537406

VERSION BM537406.1 GI:18817590

KEYWORDS EST.

SOURCE Canis familiaris (dog)

ORGANISM Canis familiaris

REFERENCE 1 (bases 1 to 529)

AUTHORS O'Shaughnessy,A.L., McCombie,W.R., Baker,J.P., Balija,V.,

Cunniss,D., Dedhia,N.N., de la Bastide,W., Katzenberger,F.,

King,L., Kirchoff,K.A., Miller,B., Muller,S., Nascimento,L.U.,

Palmer,L., Santos,L., Shah,R.S., Spiegel,L.A., Zutavern,T.,

Preston,R. and Hannon,G.J.

Expressed sequence tags from Canis familiaris (dog) (2002)

Unpublished (2002)

CONTACT: W. Richard McCombie

Lita Aurenberg Hazen Genome Sequencing Center

Cold Spring Harbor Laboratory

PO Box 100, Cold Spring Harbor, NY 11724, USA

Tel: 516 367 8884

Fax: 516 367 8874

Email: mcombie@cshl.org

Plate: ha82 row: a column: 01

Seq primer: -21M13UnivRev

High quality sequence stop: 529.

## FEATURES

## source

1..529  
 Location/Qualifiers  
 /organism="Canis familiaris"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9615"  
 /clone="ha82a01"  
 /tissue\_type="testes"  
 /note="Lib=Canis cDNAs from testes cells"  
 /notes="Vector: lambda Zap II; The library was produced by Greg Hannon and Raymond Preston (Cold Spring Harbor Laboratory). This library is oligo(dT) primed using stragene zap cDNA synthesis kit. It was made from dog testes. Please contact Greg Hannon (hannon@cshl.org) with any library related inquiries."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,178-44 Length: 529  
 Score: 68.00 Matches: 114  
 Percent Similarity: 98.28% Conservative: 0  
 Best Local Similarity: 98.28% Mismatches: 1  
 Query Match: 4.76% Indels: 2  
 DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM537406 (1-529)

QY 1037 TyrProGlnProLeuLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056  
 Db 128 TACCACAGCCCTTCTCAATCCACATGTGCTGGAGCAGCAGATACACTGCTCT 187  
 QY 1057 PheIleHisCysProLysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076  
 Db 188 TTCATTCAATGTCCAGTTCAGTCAATGGCCACAGCAAAATGCAATTAATCAGATATC 247  
 QY 1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096  
 Db 248 TGTGAGAGTGTGAGAACCTGACCACAGCAGCACTGCGACCTGCATATCTGGCTTC 307  
 QY 1097 TyrGlyAspProThrAsnGlyGlyLys-CysGlnProCysLysCysAsnGlyHisAlaSe 1116  
 Db 308 TATGGTGATCCCAACCAATGG-GGGAAGATGTGAGCCATGCAAGTGCATGGCATGCCTC 366  
 QY 1116 rLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAsnG 1136  
 Db 367 TCCTCTGTAAACCAACACGGGCAAGTCTTGTGACCACCAAGGGCGTCAAGGGGACCA 426  
 QY 1136 uCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeu 1151  
 Db 427 GTCCAGCTGTGTGAGTGGAAATCGATACCAAGGGAATCTCTC 472

## RESULT 74

BM569805/c

LOCUS

DEFINITION 530 bp mRNA linear EST 24-OCT-2003

BX669805 Sus Scrofa library (scac) Sus scrofa cDNA clone

scac00291.b.13 3prim, mRNA sequence.

ACCESSION BX669805

VERSION BX669805.1 GI:37979704

KEYWORDS EST.

SOURCE Sus scrofa

ORGANISM Sus scrofa

REFERENCE 1 (bases 1 to 530)

AUTHORS Bonnet,A., Tosser-Klopp,G., Benne,P., Cabau,C., Villegier,S.,

Soares,M., Bonaldo,F. and Hately,F.

A Pig Normalised Multi-Tissue cDNA Library